On Games of Strategic Experimentation

Dinah Rosenberg, Antoine Salomon, Nicolas Vieille

On Games of Strategic Experimentation

Dinah Rosenberg*, Antoine Salomon† and Nicolas Vieille‡

December 15, 2010

Abstract

We focus on two-player, two-armed bandit games. We analyze the joint effect on the informational spillovers between the players of the correlation between the risky arms, and the extent to which one's experimentation results are publicly disclosed. Our main results only depend on whethert informational shocks bring good or bad news. In the latter case, there is a sense in which the marginal value of these informational spillovers is zero.

Strategic experimentation issues are prevalent in most situations of social learning. In such setups, an agent may learn useful information by experimenting himself, or possibly, by observing other agents. Typical applications include dynamic R&D (see e.g. Moscarini and Squintani (2010), Malueg and Tsutsui (1997)), competition in an uncertain environment (MacLennan (1984)), financial contracting (Bergemann and Hege (2005)), etc. We refer to Bergemann and Valimaki (2008) for a recent overview of these applications.

Consider two pharmaceutical firms pursuing research on two related molecules. This research may eventually yield positive results, if the molecule or one of its derivatives turns out to have the desired medical effect, or may never yield such results if this is not the case. In the light of this uncertainty, how long should each firm be willing to wait in the absence of any positive results ? The answer is likely to depend on the informational spillovers between the two firms, on at least two grounds. On the one hand, the correlation between the two research outcomes affects the statistical inferences of the firms. In the above instance, or in, e.g., the case of two firms drilling nearby oil tracts, it is often natural to assume that these outcomes are positively correlated. Klein and Rady (2010) discuss examples from economics of law, in which the relevant outcomes are instead typically negatively correlated. On the other hand, the answer depends on the extent to which research outcomes within one firm are observed by the other firm. There are

*Economics and Decision Sciences Department, HEC Paris, and GREG-HEC; . e-mail: rosenberg@hec.fr
†Imagine, Ecole des Ponts ParisTech / CSTB, Universit Paris Est. e-mail: salomona@imagine.enpc.fr
‡Economics and Decision Sciences Department, HEC Paris, and GREG-HEC. e-mail: vieille@hec.fr.

cases in which research successes are immediately publicized (say, in the form of specific patents), and other cases in which only research policy decisions are publicly observed.

Our purpose here is to study how these two dimensions – correlation and observability – combine and influence the equilibrium outcomes. We use the continuous-time, game-theoretic model of Poisson/exponential bandits, popularized by Keller, Rady and Cripps (2005), and later studied in Keller and Rady (2010), Klein and Rady (2010), Murto and Valimaki (2009) among others. Each player is facing a two-arm bandit problem: at each time instant, each player has to choose whether to stick to a risky arm, or to switch to a safe alternative. The risky arm may be of a bad type, in which case it never delivers any payoff, or of a good type, in which case it starts delivering payoffs at a random time. In contrast with most earlier work, we assume that the decision to switch from the risky arm to the safe one is irreversible. This simplifying assumption allows us to analyze and contrast two scenarios. In one of them, both past decisions and past outcomes are publicly observed; in the second scenario, only past decisions are observed.

The basic strategic insights differ in the two scenarios. Indeed, when both actions and payoffs are observed, there is no asymmetric information between the players, and our game of timing is essentially a game of pure coordination: beyond a certain time, a player is willing to experiment only to the extent that the other player still experiments. Accordingly, the formal analysis is rather straightforward. The probability assigned by a player to his arm being good declines continuously as long as no success occurs. If the two arms are *positively* correlated, a success hit by the other player brings *good* news, and the absence of such a success is *bad* news, so that the beliefs decline faster than in a one-player setup. If the arms are *negatively* correlated, such an event instead bring *bad* news, and the belief held by the other player then jumps downwards.

When instead only actions are observed, the game does not exhibit clear strategic complementarities. In any time interval, a player deduces valuable information from the other player's behavior, only inasmuch as that player may choose to drop out in that time interval. There, an event occurs when the other player drops, and this is evidence that he did not get any payoff from the risky arm. The effect of correlation is reversed, when compared to the previous scenario. If the two arms are *positively* correlated, such an event is *bad* news. A contrario, such an event is good news if the two arms are negatively correlated.

Our main qualitative finding only hinges on whether events bring good or bad news. When events bring good news,[1] the belief threshold at which a player chooses to drop out is *the same as* the belief threshold at which he would optimally choose to drop out, if he were alone. In other words, the (marginal) option value of observing the other player is then equal to zero. In a sense, equilibria thus exhibit no encouragement effect (see Bolton and Harris (1999)). When events instead bring good news, symmetric equilibria always exhibit an encouragement effect,

[1]Thus, this is the case as soon as arms are negatively correlated and all information is public, or arms are positively correlated, and only actions are publicly observed.

and players keep experimenting with beliefs below the one-player threshold.[2]

The proofs for the two observation scenarios are fairly different. Yet, there is a common intuition. If a player is indifferent between dropping out at times t and $t + dt$, then it must be that the informational benefits derived from experimenting between t and $t + dt$ just offset the opportunity cost of doing so: thus, it might be that an event occurs, that will trigger him to continue beyond $t + dt$. But if events bring bad news, an event coming from the other player will increase that player's desire to drop out. Thus, the trade-off faced at time t is the same as if one were alone.

When payoffs (and actions) are observed, symmetric equilibria are all pure. If only actions are observed, the unique symmetric equilibrium is mixed, with a continuous distribution, and the cdf of the strategy is obtained as the solution to an linear, second-order differential equation. This difference in equilibrium structure can be understood as follows. At a symmetric equilibrium, there exists no date at which a player expects to 'learn something substantial'.[3] Otherwise indeed, the marginal value of the information obtained at that date would be positive, and this player would rather wait and get this information, rather than, say, drop out shortly before – this would contradict the equilibrium property. When payoffs are observed, a player 'learns something' from the other when this other player hits a success. This happens according to a continuous distribution. When instead only actions are observed, player i 'learns something' from player j when player j drops out. Any atom in the strategy of player j corresponds to a date at which player i's belief will jump for sure – which cannot happen at a symmetric equilibrium, as we argued. In a sense, the continuity of player j's strategy ensures that the information flow from player j to player i is continuous.

There is no asymmetric equilibrium outcome, unless when arms are negatively correlated and past payoffs are not publicly observed. In that case, there is a flury of such equilibria, and all of them weakly Pareto dominate the symmetric one. All asymmetric equilibria are purely atomic, and the two players never drop out at the same time. Yet, all equilibria share our main qualitative property: when dropping out, the belief held by a player is equal to the belief at which he would drop out in a one-player setup.

Our paper is closely linked to the literature on strategic experimentation, as we indicated above. In the Poisson/exponential setup, Keller, Rady and Cripps (2005), and Klein and Rady (2010) assume that the two risky arms are perfectly correlated (positively in the former case, and negatively in the latter) and analyze the case where payoffs are observed. However, agents may freely switch between the two arms, and these papers are thus not directly comparable to

[2] In each of the four combinations but one, there is a unique symmetric equilibrium.
[3] Formally, there exists no date at which one's belief jumps with probability one.

ours. Keller and Rady (2010) study the case where the two risks are perfectly correlated and payoffs observed, and relax the assumption that only a good risky arm may yield payoffs.

To our knowledge, existing work on the case with unobserved payoffs always assumes that the decision to switch to the safe arm is irreversible. Under the assumption that arms are perfectly (positively) correlated Rosenberg et al. (2007) provide qualitative results on the equibrium structure. The present paper is more closely related to Murto and Valimaki (2009). Within a discrete-time framework, they analyze the case in which actions are observed, and assume that risky arms are positively correlated (so that events bring bad news). Their focus is on information aggregation properties at equilibrium, especially in the limit as the number of players goes to $+\infty$. Yet, they argue that, at the symmetric equilibrium, in the continuous-time limit, a property similar to our qualitative observation does hold.

The paper is organized as follows. We lay down the model in Section 1, and state our results in Section 2. Section 3 contains proofs for the observed payoffs case, and an heuristic discussion of the observed actions case. Detailed proofs are in the Appendix.

1 The model

1.1 Setup and comments

1.1.1 The setup

Time is continuous. Each of two players is facing a strategic experimentation problem, modeled as a two-arm bandit problem. At each point in time, each player has to choose which of two arms to pull. One of the arms does not involve any uncertainty, and yields a constant payoff flow with present value s. The other arm's type is *ex ante* unknown, and it may be either *Good* or *Bad*. The decision to switch from the risky arm to the safe one is irreversible. That is, each player has to choose when, if ever, to drop out and stop experimenting. We will throughout denote by θ_i the time at which player i chooses to drop out.

A risky arm of type *Bad* (B) never yields any payoff. An arm of type *Good* (G) yields a constant payoff flow with present value γ, starting from some random time τ, but no payoff prior to τ. The random time τ follows an exponential distribution.

The players have the same prior q over the pair (R_1, R_2) of the types of the risky arms. Conditional on the type profile (R_1, R_2), the two risky arms are independent. That is, in the event where both arms are good, the two *success* times τ_1 and τ_2 associated with the risky arms of the two players are independent random variables.

We focus on symmetric problems. That is, we assume that (i) the *ex ante* probability of one's arm being good is the same for both players, (ii) when good, both risky arms have the same characteristics (values of γ and λ) and (iii) players have the same discount rate, $r > 0$. We make no assumption on the prior q beyond symmetry, and we measure correlation between the two arms by $\rho := q(R_1 = R_2 = G) - q(R_1 = G)q(R_2 = G)$. Thus, the two risky arms are positively correlated if $\rho > 0$, and negatively correlated if $\rho < 0$.

There is a very simple formal setup fitting this description. Let X_1 and X_2 be two independent random variables with an exponential distribution with parameter λ. Nature chooses the pair (R_1, R_2) of risky types according to the common prior $q \in \Delta(\{G, B\} \times \{G, B\})$. The time τ_i at which player i's risky arm starts delivering payoffs is $\tau_i := +\infty$ if $R_i = B$, and $\tau_i := X_i$ if $R_i = G$.

We will analyze two variants of this model, which differ in the information made available to the players along the play. In both variants, each player i knows at time t (i) his own past payoffs, that is, whether $\tau_i < t$ or not, and (ii) player j's past actions, that is, whether (and when) player j already dropped out, or not. In one of the two variants, no additional information is provided. In the other variant, player i observes moreover player j's past payoffs – that is, whether $\tau_j < \min(\theta_j, t)$ or not.

In the latter scenario, actions *and* payoffs are thus publicly disclosed. Since all information is public, the players always share a common posterior belief over the two types. In this scenario, the game exhibits clear strategic complementarities: the longer player j is willing to experiment, the higher the option value of experimenting for player i and, therefore, the longer player i is willing to experiment. We refer to this scenario, as the (*observed*) *payoffs* scenario.

In the former scenario, only the choices of player j are known to player i, and we refer to it as the (*observed*) *actions* scenario. At time t, the inference made by player i is contingent on player j's strategy: over any time interval $[\underline{t}, \bar{t}]$, the 'amount' of information derived by player i from player j depends on how likely it is that player j would drop out between $\underline{t}$ and $\bar{t}$ in the absence of a success.

In this scenario, there is no (low-dimensional) state variable. A complete description of a player's beliefs involves the infinite hierarchy of beliefs, and the information interaction is here more complex. As a simple illustration, observe that the optimal payoff of player i is the same, whether player j plans to drop out at time 0, or whether player j plans never to drop out, since in both cases, player i infers no information whatsoever from observing player j. This suggests that, in contrast to the payoffs scenario, player i's optimal payoff is non-monotonic in the amount of experimentation performed by player j.

1.1.2 The one-player benchmark $\mathcal{P}$

One option available to player i is to ignore altogether the information coming from player j. Thus, a natural benchmark is the decision problem $\mathcal{P}$ in which player i only observes his own payoffs. The problem $\mathcal{P}$ is a standard one-player decision problem . As long as $\tau_i \geq t$, the probability assigned by player i to his risky arm being good, decreases continuously. The unique optimal policy (see Presman (1990) or Presman and Sonin (1990) is to drop out as soon as this belief reaches

$$p_* := \frac{rs}{\lambda(\gamma - s)},$$

which happens (if no payoff is received) at the time T_p defined by $\mathbf{P}(R_i = G \mid \tau_i \geq T_p) = p_*$. As a function of his initial belief $p \in (0,1]$, the optimal payoff of the decision maker in $\mathcal{P}$ is given by

$$W(p) := gp + (s - gp_*)\frac{1-p}{1-p_*}\left(\frac{(1-p)p_*}{p(1-p_*)}\right)^{r/\lambda} \quad \text{if } p \geq p_* \text{ and } W(p) = s \text{ if } p < p_*,$$

where $g = \int_0^{+\infty} \lambda e^{-\lambda x} e^{-rx} \gamma dx = \gamma \frac{\lambda}{r+\lambda}$ is the *ex ante* expected payoff of a good risky arm. To avoid trivialities, we assume $g > s > 0$.

Thus, at equilibrium, no player ever drops out as long as he assigns a probability higher than p_* to his arm being good.

For future reference, we also define $\tilde{\mathcal{P}}$ as the decision problem in which player i observes at any time t whether $\tau_i < t$ and/or $\tau_j < t$. Equivalently, $\tilde{\mathcal{P}}$ is the best-reply problem faced by player i in the payoffs scenario, when player j plans never to drop out.

1.1.3 Good vs. Bad News

The two cases of positive vs. negative correlation are qualitatively different. Consider first the payoffs scenario. As long as players are active, and that no payoff is received, the probability assigned to one's arm being good decreases continuously. It decreases faster than in the benchmark $\mathcal{P}$ if $\rho > 0$, and slowlier if $\rho < 0$. Indeed, in the former case, (bad) news coming from the other player constitute further statistical evidence that one's arm is bad. When the first success occurs, say at time $\tau_1 < \tau_2$, player 1's arm reveals itself to be good, and player 2 updates his belief.[4] If $\rho > 0$, player 1's success is good news, and player 2's belief jumps upwards. If $\rho < 0$, this is bad news, and player 2's belief jumps downwards.

The statistical inferences in the actions variant are different. As long as both players are active, and as time goes by, the fact that player j does not drop out becomes stronger evidence that player j may have received a payoff from his risky arm. (However, the actual belief held by player i depends upon player j's strategy.) Player i's belief decreases slowlier than in $\mathcal{P}$ if $\rho > 0$, and faster if $\rho < 0$. If player j drops out, this provides conclusive evidence that $\tau_j \geq \theta_j$. This is

[4]Throughout, and depending on the context, we use the word *belief* to denote either the posterior distribution of the pair of types, or the posterior probability assigned to one's arm being good. We hope that no confusion will arise.

good news for player i if $\rho < 0$, and bad news if $\rho > 0$: player i's belief jumps upwards in the former case, and downwards in the latter case.

Both scenarios thus lead to informational shocks. Whether these shocks are interpreted as good or as bad news depends both on the sign of the correlation ρ, and on the informational scenario. As we will see, our main conclusions are mostly driven by shocks bringing good or bad news.

1.2 Strategies and outcomes

Since time is continuous, some care is needed when defining strategies. We adopt the viewpoint that a strategy dictates when to stop, as a function of all available information. We argue here that in most cases, the continuation game reduces to a one-player problem. This allows for a straightforward definition of strategies, and of expected payoffs. We treat the two scenarios in turn.

1.2.1 The payoffs scenario

Consider any time instant t, and let us focus on player i, who we assume is active prior to time t. If $\tau_i < t$, player i's risky arm is known to be good, and player i should stick to it. On the other hand, if either $\tau_j < t$ or $\theta_j < t$, player j's future behavior does not prove informative to player i. That is, player i should update his belief at time $\min(\tau_j, \theta_j)$, and proceed with the optimal policy in the one-player decision problem $\mathcal{P}$. Hence, player i's optimal decision at time t is unambiguous unless if $\min(\tau_i, \tau_j, \theta_j) \geq t$. Accordingly, we define a pure strategy of player i to be a time $t_i \in [0, +\infty]$, with the interpretation that player i drops out at time t_i *if* player j has not dropped out, and *if* no payoff has been received prior to t. A mixed strategy is a probability distribution over $[0, +\infty]$.

We now explicit how to compute the expected payoff $\gamma^i_P(t_1, t_2)$ induced by an arbitrary pure profile (t_1, t_2).

Assume first that $t_1 = t_2 = t$. The outcome of the game is determined as follows. In the event where $\min(\tau_1, \tau_2) \geq t$, both players drop out at time t. If a success occurs prior to t, say at time $\tau_1 < \min(\tau_2, t)$, player 2 updates his belief to $\psi(\tau_1)$, where $\psi(x) := \mathbf{P}(R_2 = G \mid R_1 = G, \tau_2 \geq x)$. Player 2 then remains active until his belief reaches p_*, with an expected continuation payoff equal to $W(\psi(\tau_1))$. Thus, when $t_1 = t_2$, the expected payoff of player i is

$$\gamma^i_P(t_i, t_j) := \gamma \mathbf{E}\left[e^{-r\tau_i} 1_{\tau_i < \min(\tau_j, t_i)}\right] + \mathbf{E}\left[e^{-r\tau_j} W(\psi(\tau_j)) 1_{\tau_j < \min(\tau_i, t_i)}\right] + s e^{-rt_i} \mathbf{P}(\min(\tau_i, \tau_j) \geq t_i)$$

(here and in the sequel, we use without further notice the fact that the probability that τ_1 and τ_2 coincide and be finite is equal to zero: $\mathbf{P}(\tau_1 = \tau_2 < +\infty) = 0$).

The expected payoff $\gamma_P^i(t_i,t_j)$ can alternatively be computed as follows. Conditional on $R^i=B$, player i's overall payoff is zero. Next, conditional on $R^i=G$, and $R^j=B$, player i's overall payoff is $e^{-r\tau_i}\gamma$, provided that $\tau_i<t_i$. The conditional expected payoff (given $R^i=G, R^j=B$) is thus equal to $\mathcal{A}:=\int_0^{t_i}\lambda e^{-\lambda x}e^{-rx}dx=\frac{\lambda}{r+\lambda}\gamma(1-e^{-(r+\lambda)t_i})$. Finally, conditional on $R^i=R^j=G$, player i's overall payoff is equal to $e^{-r\tau_i}\gamma$, provided that either $\tau_i<t_i$, or that $\tau_j<t_i$ and $\tau_i\in[t_i,T]$, where T solves $\psi(T)=p_*$. The conditional expected payoff (given $R^i=R^j=G$) is thus equal to

$$\mathcal{B}:=\frac{\lambda}{r+\lambda}\gamma(1-e^{-(r+\lambda)t_i})+(1-e^{-\lambda t_i})\gamma\int_{t_i}^{T}\lambda e^{-\lambda x}e^{-rx}dx.$$

Then, $\gamma_P^i(t_i,t_j)$ is equal to the weighted sum of $\mathcal{A}$ and $\mathcal{B}$, using the prior q as weights.

Assume next that $t_1<t_2$. In the event where $\min(\tau_1,\tau_2)\geq t_1$, player 1 drops out at time t_1, and player 2 next waits until his belief reaches p_*. Player 2's continuation payoff at time t_1 is given by $W(\phi(t_1))$, where $\phi(x):=\mathbf{P}(R_2=G\mid \tau_1\geq x,\tau_2\geq x)$ is player 2's belief at time x if no success occurred prior to x. If instead $\min(\tau_1,\tau_2)<t_1$, the 'non-successful' player updates his belief at time $\min(\tau_1,\tau_2)$ to $\psi(\min(\tau_1,\tau_2))$, and again, switches to the optimal policy in $\mathcal{P}$, with a continuation payoff equal to $W(\psi(\min(\tau_1,\tau_2))$.

We stress that the outcome of the profile (t_1,t_2) does not depend on t_2 as long as $t_2>t_1$. Observe also that player 1's expected payoff does not depend on t_2, as long as $t_2\geq t_1$. Thus, $\gamma_P^1(t_1,t_2)=\gamma_P^1(t_1,t_1)$, while

$$\gamma_P^2(t_1,t_2)=\gamma\mathbf{E}\left[e^{-r\tau_2}1_{\tau_2<\min(\tau_1,t_1)}\right]+\mathbf{E}\left[e^{-r\tau_1}W(\psi(\tau_1))1_{\tau_1<\min(\tau_2,t_1)}\right]+e^{-rt_1}W(\phi(t_1))\mathbf{P}(\min(\tau_1,\tau_2)\geq t_1).$$

1.2.2 The actions scenario

We now turn to the definition of strategies and to the computation of payoffs in the actions scenario. Again, consider any time instant $t\in\mathbf{R}^+$, and let us focus on player i, who we assume remained active prior to t. As above, if $\tau_i<t$, player i's arm is known to be good, and player i therefore sticks to it. If now $\theta_j<t$,[5] this is evidence that $\tau_j\geq\theta_j$. At time θ_j, player i updates his belief to $\phi(\theta_j)$ $(=\mathbf{P}(R_i=G\mid\min(\tau_i,\tau_j)\geq\theta_j))$, stays alone and therefore remains active until his belief reaches p_*. Hence, player i's optimal decision at t is dictated by $\mathcal{P}$, unless if $\min(\tau_i,\theta_j)\geq t$. Here again, we thus let a pure strategy of player i be a time $t_i\in[0,+\infty]$, with the interpretation that player i drops out at t_i if $\tau_i\geq t_i$ and $\theta_j\geq t_i$. Again, a mixed strategy is a probability distribution over $[0,+\infty]$.

Even though a pure strategy is the same mathematical object in the two scenarios, the computation of expected payoffs is of course different. Consider an arbitrary pure profile (t_1,t_2).

[5] Recall that τ_j is not observed by player i.

If $t_1 = t_2$, player i drops out at t_i if $\tau_i \geq t_i$, and stays forever if $\tau_i < t_i$. Player i's expected payoff $\gamma_A^i(t_1, t_2)$ is thus given by

$$\gamma_A^i(t_1, t_2) := \gamma \mathbf{E}\left[e^{-r\tau_i} 1_{\tau_i < t_i}\right].$$

Assume instead that $t_1 < t_2$. Player 1's behavior is as we just described: he drops out at time t_1 if $\tau_1 \geq t_1$, and stays forever otherwise. Hence, player 1's payoff is given by

$$\gamma_A^1(t_1, t_2) = \gamma \mathbf{E}\left[e^{-r\tau_1} 1_{\tau_1 < t_1}\right].$$

The payoff of player 2 is to be computed as follows. If player 1 drops out at time t_1, player 2 updates his belief to $\phi(t_1)$, and proceeds with the optimal policy in $\mathcal{P}$. If player 1 does not drop out at t_1, player 2 either drops out at t_2, or never, according to whether $\tau_2 \geq t_2$ or $\tau_2 < t_2$. Hence, player 2's payoff is given by

$$\gamma_A^2(t_1, t_2) := e^{-rt_1} W(\phi(t_1)) \mathbf{P}(\min(\tau_1, \tau_2) \geq t_1) + \gamma \mathbf{E}\left[e^{-r\tau_2}\left(1_{\tau_2 < t_1} + 1_{\tau_1 < t_1, \tau_2 \in [t_1, t_2)}\right)\right].$$

As in the payoffs scenario, simple integral expressions for these payoffs will be used.

1.2.3 Conceptual issues

The definitions and interpretations of strategies raise a few conceptual issues. For concreteness, let us focus on the actions scenario, and let a pure profile (t_1, t_2) be given, with $t_1 < t_2$. According to the previous interpretation, player 2 chooses not to drop out at time t_1. Assume that the event $\theta_1 = t_1 < \tau_2$ materializes.[6] Again according to our interpretation, player 1 drops out at t_1, player 2 then updates his belief to $\phi(t_1)$, and switches to the optimal policy in $\mathcal{P}$. Depending on t_1 and on ρ, it may well be that $\phi(t_1) < p_*$, in which case player 2 wishes to stop *as early as possible* following t_1 and even, possibly, to overturn his time t_1's decision. How then should player 2's decision at time t_1 be defined ? Preventing player 2 to drop out at time t_1 after he has seen that player 1 has exited, would lead to spurious equilibrium non-existence results. As a matter of fact, expected payoffs would not even be defined.[7]

To avoid this, we implicitly adopted the view that player 2 is allowed to revise his time t_1 decision, should player 1 decide to drop out. This leads to the informal idea that each player has two decision nodes at each time instant t. At the first one, a player conditions his decision on information available *prior to t*. Should he decide to remain active, and should some informational event occur at time t, he is next allowed to revise his decision – that is, to drop

[6] Given the players's strategies, this is the positive probability event where no payoff has been received prior to t_1.

[7] Similar issues arise in the payoffs scenario as well.

out –, with *no* delay.[8] We leave aside the technical issues involved in defining a game tree that fits with these interpretations. However, the previous formulas for expected payoffs lead to a well-defined game in strategic form, in which the space of pure strategies is $\mathbf{R}_+ \cup \{+\infty\}$.

We emphasize that a strategy only dictates when to stop. That is, the pure strategy t does not specify when to stop, in the counterfactual event where one would not stop at time t when required. This prevents us from discussing equilibrium refinements, and all of our results are stated in terms of Nash equilibrium.

2 Results

As noted, our qualitative results hinge on the nature of informational shocks – whether they bring good news or bad news. Consequently, we organize the discussion of our results around this criterion.

Three belief functions, ϕ, ψ and ξ play an important role in the analysis. The first two have already been introduced, and we here recall their definitions. The functions ϕ and ψ are defined as

$$\phi(t) := \mathbf{P}(R^i = G \mid \tau_i \geq t, \tau_j \geq t) \text{ and } \psi(t) := \mathbf{P}(R^i = G \mid R^j = G, \tau_i \geq t).$$

Thus, $\phi(t)$ is the belief held by a player at time t conditional on no success occurring prior to t, while $\psi(t)$ is belief of player i if he knows player j's arm to be good. Note that both ϕ and ψ are continuous and decreasing.[9]

We also set $p(t) := \mathbf{P}(R^i = G \mid \tau_i \geq t)$. Note that $\psi(t) > p(t) > \phi(t)$ if $\rho > 0$, while $\psi(t) < p(t) < \phi(t)$ if $\rho < 0$.

Given a player j's strategy, the function ξ_i is defined as

$$\xi_i(t) := \mathbf{P}(R^i = G \mid \theta_j \geq t, \tau_i \geq t).$$

We stress that $\xi_i(t)$ depends on player j's strategy, although this does not appear in the notation. In particular, the continuity and monotonicity properties of ξ_i are contingent on player j's strategy.

One can check that ξ_i is left-continuous, and has a right-limit at each $t \in \mathbf{R}_+$, which is given by $\xi_i(t+) = \mathbf{P}(R^i = G \mid \theta_j > t, \tau_i > t)$. In addition, the map ξ_i is decreasing if $\rho < 0$, irrespective of the strategy of player j.

[8]Obviously, one cannot merge these two nodes into a single decision node, at which each player would be allowed to condition his time t-decision on the other player time t-decision. Indeed, circularity problems would arise, leading to the indeterminacy of the outcome.

[9]Unless if risky arms are perfectly negatively correlated, in which case ϕ is constant. In what follows, we sometimes implicitly assume that arms are not perfectly negatively correlated, but all of our results cover that case as well.

2.1 Negative informational shocks

We start by discussing here the case where events bring bad news. This is the case when either (i) $\rho < 0$ in the payoffs scenario, or (ii) $\rho > 0$ in the actions scenario.

2.1.1 Main result

We let dates T_p, T_ϕ and T_ψ be defined by

$$p(T_p) = p_*, \ \phi(T_\phi) = p_* \text{ and } \psi(T_\psi) = p_*.$$

Theorem 1 (payoff scenario, $\rho < 0$) *There is a unique symmetric equilibrium, (T_ϕ, T_ϕ).*

Theorem 2 (actions scenario, $\rho > 0$) *There is a unique symmetric equilibrium, (σ_*, σ_*). The distribution σ_* has a support*[10] *$[T_p, T_\psi]$, and a density, which is positive and continuous on $[T_p, T_\psi]$.*

One has $\xi_i(t) = p_*$ for each $t \in [T_p, T_\psi]$.

The cdf of σ_* is the unique solution of a linear differential equation, see Appendix.

The main conclusion to be drawn is that, at equilibrium, no player is willing to keep experimenting with a belief below p_*, the optimal cut-off in $\mathcal{P}$. Put otherwise, while the informational spill-over/externality clearly affects the speed at which beliefs change, the marginal option value of observing the other player is equal to *zero*, when informational events bring bad news. Indeed, when player i's belief reaches p_*, he finds it optimal to exit, *whether or not* he may benefit from observing the other player's action/payoffs in the future.

While the equilibrium uniqueness claim in Theorems 1 and 2 is partly an artefact of the symmetry requirement, this main conclusion holds at all asymmetric equilibria as well, as we will see below.

There is a simple, yet very partial, intuition behind Theorems 1 and 2. Assume that, at some time $t \in \mathbf{R}^+$, player i is considering whether to drop out at time t, or to wait an additional, infinitesimal amount of time dt before dropping out. The rationale for waiting an additional dt is that there is a chance that some event will occur that will cause a discontinuity in player i's belief, and trigger him to continue beyond $t + dt$. Such an event may either be one's own success, or an event coming from the observation of player j.[11] In the latter case however, that event would bring bad news, and would only reinforce player i's desire to drop out. Hence, only the prospect of hitting a success may justify to wait. But then, the condition of being indifferent between exiting at time t or at time $t + dt$ writes the same way, whether in the decision problem $\mathcal{P}$ or when player i is observing player j. This sketch abstracts away from a number of complications, and the proofs of Theorems 1 and 2 differ significantly.

[10] That is, the smallest closed set in $[0, +\infty]$ that is assigned probability 1 by σ_*.

[11] That is, player j getting a payoff, or dropping out, depending on the scenario.

2.1.2 Beyond the symmetry requirement

We here provide evidence that our main conclusion does not rely on our restriction to symmetric equilibria. We discuss first the payoffs scenario (with $\rho < 0$). Note first that $\phi(t) > p_*$ for $t < T_\phi$, hence no player ever stops prior to T_ϕ at equilibrium. We will actually prove that the unique optimal strategy in the decision problem[12] $\tilde{\mathcal{P}}$ is to exit at T_ϕ. This will imply that there is a unique equilibrium outcome, in which each player drops out at T_ϕ in the absence of payoffs.

Proposition 1 (payoff scenario, $\rho < 0$) *The equilibria are the profiles (T_ϕ, σ), where σ is any distribution such that $\sigma([T_\phi, +\infty]) = 1$, together with the profiles obtained when exchanging the two players.*

Since the outcomes induced by (T_ϕ, T_ϕ) and (T_ϕ, σ) are the same whenever $\sigma([T_\phi, +\infty]) = 1$, this multiplicity of equilibria is spurious.

In the actions scenario (with $\rho > 0$), we do not have a complete characterization of the set of equilibria and, as Proposition 3 below shows, there may be countably many different equilibrium outcomes. However, all the equilibria share the main qualitative feature of the symmetric one.

Proposition 2 (actions scenario, $\rho > 0$) *Let (σ_1, σ_2) be an equilibrium, and let $S_i \subseteq \mathbf{R}_+ \cup \{+\infty\}$ denote the support of σ_i. One has $\xi_i(t) = \xi_i(t+) = p_*$ for each $t \in S_i$ $(i = 1, 2)$.*

Proposition 2 can be interpreted as follows. Consider any date t at which player i may potentially drop out, according to his equilibrium strategy σ_i. Then, the belief held by player i, whether computed before observing player j's decision at date t ($\xi_i(t)$) or after observing it ($\xi_i(t+)$) is equal to the one-player optimal threshold, p_*.

According to Proposition 3 below, for any $k \in \mathbf{N}$, there is an (asymmetric) equilibrium (σ_1, σ_2) in which each of the two strategies assigns positive probability to exactly k dates.

Proposition 3 (actions scenario, $\rho > 0$) *For some specification of the parameters, the following statement is true. For every $k \in \mathbf{N}$, there exist two sets $\mathcal{U} = \{u_1, \ldots, u_k\} \subset \mathbf{R}_+$ and $\mathcal{V} = \{v_1, \ldots, v_k\} \subset \mathbf{R}_+$ and an equilibrium (σ_1, σ_2) such that*

- $T_p = u_1 < v_1 < \cdots < u_k < v_k = T_\psi$;
- *The supports of σ_1 and σ_2 are respectively equal to $\mathcal{U}$ and $\mathcal{V}$.*

We believe that the conclusion actually holds true for all parameter values, but leave this outside of the scope of this paper.

[12] Recall that $\tilde{\mathcal{P}}$ is the benchmark decision problem in which the decision maker observes, on top of his own risky arm, whether the other risky arm produces any payoff.

2.2 Positive informational shocks

We now present our results for the case where events bring good news. This is the case when either (i) $\rho > 0$ in the payoffs scenario or (ii) $\rho < 0$ in the actions scenario.

In the payoffs scenario, the game is a game of pure coordination. Player i is willing to experiment beyond T_ϕ only to the extent that player j also experiments beyond T_ϕ. On the other hand, no player will ever be willing to experiment beyond T_{**}, the (unique) optimal policy in the decision problem $\tilde{\mathcal{P}}$.

As we will show in the Appendix, the optimal strategy in $\tilde{\mathcal{P}}$ is $T_{**} = T_\phi$ if $\rho < 0$. That is, the optimal belief threshold in the decision problem $\tilde{\mathcal{P}}$ is then the same as the optimal belief threshold in $\mathcal{P}$. This explains why there is a unique equilibrium in that case (see Theorem 1 in the previous section).

When $\rho > 0$, one has $T_{**} > T_\phi$ and, not surprisingly, a continuum of equilibrium outcomes.

Theorem 3 (payoff scenario, $\rho > 0$) *One has $T_{**} > T_\phi$. All symmetric equilibria are pure, and all pure equilibria are symmetric. The symmetric equilibria are the pairs (T, T), with $T \in [T_\phi, T_{**}]$.*

The asymmetric equilibria are the profiles (T_{**}, σ), where σ assigns probability 1 to $[T_{**}, +\infty]$ (together with the profiles obtained when exchanging the two players).

Observe that in all equilibria (with the exception of the equilibrium (T_ϕ, T_ϕ)), players remain active with beliefs below p_*.

Conclusions are much sharper in the actions scenario. We introduce the first-order, integro-differential equation

$$\frac{W(\phi(t)) - s}{\lambda(\gamma - s)} F'(t) = (\phi(t) - p_*) F(t) + \lambda \phi(t)(\psi(t) - p_*) \int_0^t e^{\lambda(t-x)} F(x) dx. \tag{1}$$

Standard results on differential equations imply that there is a unique function, which we denote F_*, that solves (1) on the interval $[T_p, T_\phi)$, and such that $F_*(t) = 0$ for all $t \le T_p$.

Theorem 4 (actions scenario, $\rho < 0$) *There is a unique equilibrium, which is a symmetric and non-atomic equilibrium (σ, σ).*

The support of the distribution σ is equal to the interval $\left[T_p, \hat{T}\right]$, where $\hat{T} := \min\{t : F_*(t) = 1\} < T_\phi$. The distribution σ has a positive and continuous density on its support, and its cdf is equal to F_* on the interval $[T_p, T_\phi]$.

Since $\rho < 0$, one has $\xi_i(t) \le p(t) \le p_*$, and the two inequalities are strict for $t > T_p$. That is, in contrast with the case where $\rho > 0$, players remain active with beliefs below p_*.

The proof of Theorem 4 is more delicate. We here discuss some rough intuition behind the result. Beyond the fact that beliefs are strategy-contingent, there is one feature that distinguishes the actions scenario from the payoffs scenario. Assume that player j uses the pure strategy t. When facing this strategy, player i *anticipates* prior to t, that he will receive a piece of information at t. This obviously creates an incentive for player i to wait until after t, when player j's action choice will become known. This rules out the existence of a pure symmetric equilibrium and, more generally, severely limits the scope for equilibria involving atoms. The reason why equilibria involving atoms fail to exist if $\rho < 0$, while many such equilibria exist if $\rho > 0$, is discussed at length in Section 3. As a fact, the most intricate part of the proof of Theorem 4 consists in showing that the cdf's of equilibrium strategies are continuous.

Let us take this conclusion for granted, and let an equilibrium (σ_1, σ_2) be given, with continuous cdf's. For concreteness, let us focus on player 1. Observe first that player 1 keeps being more pessimistic with time, as long as $\tau_1 \geq t$ and $\theta_2 \geq t$. This is the combined effect of two factors. First, the fact that player 1 does not get any payoff is bad news in itself. Second, the fact that player 2 does not drop out is an increasingly stronger indication that player 2 may have received a payoff with the risky arm, which makes player 2 even more pessimistic since $\rho < 0$.

On the other hand, the event of player 2 exiting becomes lesser good news with time. Indeed, if player 2 drops out at time t, player 1's continuation payoff jumps to $W(\phi(t))$, and that payoff is decreasing with t.

Since, by the equilibrium condition, player 1 is indifferent between exiting or not at all times in the support of σ, it must be that the chances of observing these good news *increase* with time. That is, the hazard rate of the exit decision of player 2 should increase with time, in a way that exactly offsets the two effects identified above. This condition writes as Equation (1).

This intuition suggests that the support and the density of σ_i are uniquely dictated by the requirement that player j be indifferent between all times in the support of σ_j. It also suggests why both strategies σ_1 and σ_2 must have the same, connected, support – hence the uniqueness of the equilibrium.

3 Proofs

We find it more natural and convenient to present the payoffs scenario first. Most proofs in that case are short and technically elementary, and we thus provide full details. The only one exception is the analysis of the decision problem $\tilde{\mathcal{P}}$, for which technical details are postponed to the Appendix.

By contrast, all proofs in the actions scenario require technical care. Therefore, we will provide detailed sketches here, and will refer the reader to the Appendix for the complete analysis.

3.1 The payoffs scenario

The analysis of the payoffs scenario rests upon the analysis of the one-player decision problem $\tilde{\mathcal{P}}$, which we now provide.

3.1.1 The marginal value of observing the other player

We recall that a (pure) policy in $\tilde{\mathcal{P}}$ is a time $t \in \mathbf{R}_+ \cup \{+\infty\}$, with the understanding that the decision maker, say player 1, will choose (i) to drop out at t if $\min(\tau_1, \tau_2) \geq t$, (ii) to drop out at $t' := \min\{x \geq t : \psi(x) \leq p_*\}$ if both $\tau_2 < t \leq \tau_1$ and $\tau_1 \geq t'$, and (iii) never to drop out otherwise.

We will show in the Appendix that the decision problem $\tilde{\mathcal{P}}$ admits a unique optimal policy, $T_{**} < +\infty$. We denote by $p_{**} := \phi(T_{**})$ the belief held by the decision maker at time T_{**}, in the event where $\min(\tau_1, \tau_2) \geq T_{**}$. (Plainly, one has $p_{**} \leq p_*$, the optimal threshold in $\mathcal{P}$.)

Proposition 4 *One has $T_{**} = T_\phi$ if $\rho < 0$, and $T_{**} > T_\phi$ if $\rho > 0$.*

We here limit ourselves with an heuristic proof of Proposition 4. Let us place ourselves at time $t := T_{**}$, and assume that $\min(\tau_1, \tau_2) \geq t$. By a dynamic programming principle, the decision maker is indifferent between, on the one hand, dropping out at time t and, on the other hand, staying in for an additional, short, amount of time dt, then behaving in an optimal way at time $t + dt$, in the light of the information acquired between t and $t + dt$.

There are three events $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ to be considered at time $t + dt$:

$\mathcal{A}$: player 1 hit a success in this additional amount of time ($\tau_1 \in [t, t + dt)$),

$\mathcal{B}$: player 2 hit a success but player 1 did not ($\tau_2 \in [t, t + dt), \tau_1 \geq t + dt$),

$\mathcal{C}$: no one did.

As is usual for continuous-time dynamic programming, we write a first-order expansion of the above indifference condition.

The (conditional) probability that $\mathcal{A}$ occurs is (approximately) $\lambda p_{**} dt$. If event $\mathcal{A}$ occurs, the decision maker's overall continuation payoff, (discounted back to time t) is approximately equal to γ.[13]

Observe next that the (conditional) probability that *both* arms yield a payoff between t and $t + dt$ is of the order of $(dt)^2$, irrespective of the correlation between the two arms' types. Therefore, the (conditional) probability of $\mathcal{B}$ occurring, is approximately equal to the conditional probability that $\tau_2 \in [t, t + dt)$, which is $\lambda p_{**} dt$. If the event $\mathcal{B}$ occurs, the flow payoff between

[13] This overall payoff depends on the exact time τ_1. However, since it is multiplied by a probability which is itself of the order of dt, only the leading term has to be kept.

t and $t + dt$ is equal to zero, and the continuation payoff of the decision maker is approximately $W(\psi(t))$.

Finally, the conditional probability that $\mathcal{C}$ occurs is the residual probability, $1 - 2\lambda p_{**}dt$. On the event $\mathcal{C}$, the flow payoff is exactly 0, while the continuation payoff is *exactly* s, since the decision maker then drops out at time $t + dt$. Hence on the event $\mathcal{C}$, the overall payoff is equal to $e^{-rdt}s$, when discounted back to t.

As a consequence, the first-order expansion of the indifference condition writes

$$s = \lambda p_{**}\gamma dt + \lambda p_{**}W(\psi(T_{**}))dt + (1 - rdt)\,(1 - 2\lambda p_{**}dt)\,s,$$

which simplifies into

$$\lambda p_{**}\gamma + \lambda p_{**}W(\psi(T_{**})) - (r + 2\lambda p_{**})s = 0. \tag{2}$$

If $\rho < 0$, then $\psi(T_{**}) < \phi(T_{**}) \leq \phi(T_\phi) = p_*$, hence $W(\psi(T_{**})) = s$. In that case, equation (2) yields

$$p_{**} = \frac{rs}{\lambda(\gamma - s)} = p_*,$$

and $T_{**} = T_\phi$ as claimed.

If $\rho > 0$, one has $\psi(T_\phi) > \phi(T_\phi)$, hence $W(\psi(T_\phi)) > s$, and T_ϕ is no longer a solution of equation (2). Therefore, $T_{**} > T_\phi$.

Note that, according to equation (2), T_{**} is a solution to the equation

$$\phi(t)\,(\gamma + W(\psi(t)) - 2s) = \frac{rs}{\lambda}. \tag{3}$$

Since the left-hand side of equation (3) is the product of the positive and decreasing function ϕ with a non-increasing function of t, the optimal time T_{**} is uniquely determined by equation (3).

3.1.2 Proofs of Theorems 1 and 3

We here provide complete proofs of Theorems 1 and 3 for the payoffs scenario. We first describe the best-reply correspondence of either player.

We denote by $\tilde{\pi}$ the payoff function in the decision problem $\tilde{\mathcal{P}}$. In the light of Section 1.2.1, player i's expected payoff $\gamma^i_P(t_i, t_j)$ is equal to $\tilde{\pi}(t_i)$ as long as $t_i \leq t_j$, and does not depend on t_i whenever $t_i > t_j$.

The continuity properties of $\gamma_P(\cdot, \cdot)$ on the diagonal $t_1 = t_2$ play a role in the analysis. We fix $t_1 \in \mathbf{R}^+$, and we compare the outcomes induced by the profiles (t_1, t_1) and (t_1, t_2), where $t_2 > t_1$. The two profiles yield the same outcome, except possibly if $\min(\tau_1, \tau_2) \geq t_1$. On the latter event, both players do exit at time t_1 according to the profile (t_1, t_1). Instead, under the profile (t_1, t_2), player 1 drops out at time t_1, and player 2 remains active until his belief reaches

p_*, with a continuation payoff equal to $W(\phi(t_1)) \geq s$. Formally,

$$\gamma_P^2(t_1, t_2) - \gamma_P^2(t_1, t_1) = \mathbf{P}(\min(\tau_1, \tau_2) \geq t_1) e^{-rt_1} \left(W(\phi(t_1)) - s \right).$$

The continuation payoff $W(\phi(t_1))$ exceeds s if and only if $\phi(t_1) > p_*$, that is, iff $t_1 < T_\phi$. It follows that $\gamma_P^2(t_1, t_1) \leq \gamma_P^2(t_1, t_2)$, and $\gamma_P^2(t_1, t_1) < \gamma_P^2(t_1, t_2)$ if and only if $t_1 < T_\phi$.

This observation allows us to pin down the best reply correspondence. Fix first a *pure* strategy t_1 of player 1. The map $t_2 \mapsto \gamma_P^2(t_1, t_2)$ coincides with $\tilde{\pi}(\cdot)$ on $[0, t_1]$, and is constant on $(t_1, +\infty]$, with $\gamma_P^2(t_1, t_1+) \geq \gamma_P^2(t_1, t_1)$, and $\gamma_P^2(t_1, t_1+) > \gamma_P^2(t_1, t_1)$ if and only if $t_1 < T_\phi$. Since $\tilde{\pi}(\cdot)$ is single-peaked, with a maximum at T_{**}, it follows that the set $B_2(t_1)$ of pure best replies to t_1 is given by (i) $B_2(t_1) = (t_1, +\infty]$ if $t_1 < T_\phi$, (ii) $B_2(t_1) = [t_1, +\infty]$ if $t_1 \in [T_\phi, T_{**}]$, and (iii) $B_2(t_1) = \{T_{**}\}$ if $t_1 > T_{**}$.

In particular, if (t, t) is a pure symmetric equilibrium, then $t \geq T_\phi$ by (i) and $t \leq T_{**}$ by (ii), hence $t \in [T_\phi, T_{**}]$. Conversely, any profile (t, t) such that $t \in [T_\phi, T_{**}]$ is an equilibrium.

Much more precise conclusions are readily available. Observe that the pure strategy T_{**} is a best reply to *any* pure strategy of player 1 and, therefore, to any mixed strategy as well. Fix now a mixed strategy σ_1 of player 1, and let $t_2 \in [0, +\infty]$ be arbitrary. Let us compare the outcomes induced by the two profiles (σ_1, t_2) and (σ_1, T_{**}). On the event $\min(\tau_1, \tau_2, \theta_1) < \min(t_2, T_{**})$, the two outcomes coincide. Assume instead that the event $\min(\tau_1, \tau_2, \theta_2) \geq \min(t_2, T_{**})$ occurs. If $t_2 > T_{**}$, it is strictly suboptimal to continue any longer beyond T_{**}, hence the expected continuation payoff, conditional on $\min(\tau_1, \tau_2, \theta_1) \geq T_{**}$ is strictly higher under (σ_1, T_{**}). Thus, one has $\gamma_P^2(\sigma_1, t_2) = \gamma_P^2(\sigma_1, T_{**})$ if and only if $\sigma_1((T_{**}, +\infty]) = 0$.

If now $t_2 < T_{**}$, the only optimal policy is to continue until Player 1 drops. Thus, conditional on $\min(\tau_1, \tau_2, \theta_1) \geq t_2$, the expected continuation payoff of player 2 is strictly higher under (σ_1, T_{**}). In that case, $\gamma_P^2(\sigma_1, t_2) = \gamma_P^2(\sigma_1, T_{**})$ if and only if $\sigma_1((t_2, +\infty]) = 0$.

This allows us to characterize the set of all equilibria. Since γ_P^i is not a continuous function, a little care is needed. Let (σ_1, σ_2) be an equilibrium, and denote by S_i the support of the distribution σ_i. By the equilibrium property, one has $\gamma^i(t_i, \sigma_j) = \gamma^i(\sigma_i, \sigma_j)$ $(= \gamma^i(T_{**}, \sigma_j))$ for σ_i-a.e. $t_i \in S_i$. Set $\underline{t}_i = \min S_i$ and $\bar{t}_i = \max S_i$. Note that $\underline{t}_i \geq T_\phi$ for $i = 1, 2$.

Assume that, say, $\underline{t}_1 < T_{**}$.[14] Since $\underline{t}_1 \in S_1$, there exist strategies t_1 arbitrarily close to $\underline{t}_1$, such that $\gamma^1(t_1, \sigma_2) = \gamma^1(T_{**}, \sigma_2)$. By the previous paragraph, one has $\sigma_2((t_1, +\infty]) = 0$ for any such t_1. This implies that $\bar{t}_2 \leq \underline{t}_1$ and, *a fortiori*, $\underline{t}_2 < T_{**}$. This allows us to exchange the roles of the two players in this argument. One then obtains $\bar{t}_1 \leq \underline{t}_2$. Hence, $\underline{t}_1 = \bar{t}_1 = \underline{t}_2 = \bar{t}_2$: (σ_1, σ_2) is a pure symmetric equilibrium.

Assume now that, say, $\bar{t}_1 > T_{**}$. The previous argument shows that $\underline{t}_1, \underline{t}_2 \geq T_{**}$. Since $\bar{t}_1 \in S_1$, there exist strategies $t_1 > T_{**}$ such that $\gamma_P^1(t_1, \sigma_2) = \gamma_P^1(T_{**}, \sigma_2)$. Therefore, $\sigma_2((T_{**}, +\infty]) = 0$.

[14] If $\rho < 0$, one has $T_{**} = T_\phi$, hence this situation cannot possibly arise.

Hence, σ_2 is the pure strategy T_{**}, while σ_1 is concentrated on $[T_{**}, +\infty]$. Conversely, any such pair (σ_1, σ_2) is an equilibrium. This concludes the proof of Theorems 1 and 3, and of Proposition 1 as well.

3.2 The actions scenario

We here provide the main insights into the actions scenario, but we postpone all technical details to the Appendix. While we remain at an informal level, we follow rather closely the formal proof, for the reader's convenience. In Section 3.2.1, we gather a number of observations, which hold irrespective of the correlation. We will next discuss the cases of a positive and of a negative correlation in turn.

We recall that a (pure) strategy of player i is a time $t \in \mathbf{R}_+ \cup \{+\infty\}$, with the interpretation that player i exits at time t if $\min(\tau_i, \theta_j) \geq t$, and otherwise behaves optimally from time $\min(\theta_j, \tau_i)$ on.

For most of the section, we let a tentative equilibrium (σ_1, σ_2) be given, and we proceed using necessary equilibrium conditions. Throughout, we denote by S_i the support of the distribution σ_i.

3.2.1 Atoms, beliefs and exit

We first discuss the impact on the exit decision of player i, of atoms in the distribution σ_j.

Recall that, when at time t, player i only knows whether player j dropped *prior to* t, that is, whether $\theta_j < t$. Whether player j drops out at time t will become known immediately after t. As we argue in the Appendix, this implies that when facing σ_j, player i's payoff function $t \mapsto \gamma_A^i(t, \sigma_j)$ is left-continuous, with a right-limit, $\gamma_A^i(t+, \sigma_j)$ at each t. The belief function $t \mapsto \xi_i(t)$ enjoys similar continuity properties.

If $\sigma_j(\{t\}) = 0$, waiting an infinitesimal amount of time beyond t before possibly dropping out, makes no difference : player i's payoff function $\gamma_A^i(\cdot, \sigma_j)$ is then continuous at t, and the belief $\xi_i(\cdot)$ is continuous as well.

If instead there is a positive probability that player j drops out at t, (that is, if σ_j has an atom at t), then waiting may make a difference. Indeed, if player j choose to drop out, the continuation payoff of player i is then $W(\phi(t))$ rather than s, had player j dropped out at t. Formally, one has

$$\gamma_A^i(t+, \sigma_j) - \gamma_A^i(t, \sigma_j) = \mathbf{P}(\tau_i \geq t, \theta_j = t)\,(W(\phi(t)) - s)\,. \qquad (4)$$

Note in particular that, whenever the payoff is discontinuous, it jumps *upwards.*

The right-continuity of the payoff function readily implies the observation **O1** below.

O1 If $\gamma_A^i(t+,\sigma_j) > \gamma_A^i(t,\sigma_j)$, then $\gamma_A^i(t',\sigma_j) < \gamma_A^i(\sigma_i,\sigma_j)$ for every $t' \leq t$ close enough to t.

This observation formalizes the following intuition. Under the assumption in **O1**, the optimal continuation payoff of player i[15] at time $t+$ *does* depend on player 2's decision at time t. And since there is a positive probability that player j stops at t, the information gotten at time t has a positive value. Then, when at time $t' \leq t$ close to t, the continuation payoff of player i exceeds s if he waits until told the decision of player j at time t, and then behaves optimally. Hence it is suboptimal to exit at t'.

We next discuss the relation between the belief $\xi_i(t)$ held by player i at time t, and his optimal decision at time t.

We first formalize the intuition that, when this belief exceeds p_*, it it optimal not to drop out. Assume that $\xi_i(t+) > p_*$. Thus, when at time t, and *after* hearing player j's decision, player i is optimistic – even if he were to be alone, he would find it optimal to experiment. Again, this has a simple consequence. When at time $t+$,[16] the continuation payoff of player i is higher than s if he chooses to experiment for at least a short additional amount of time, and next to behave optimally. This leads to our second observation.

O2 If $\xi_i(t+) > p_*$, then $\gamma_A^i(t',\sigma_j) < \gamma_A^i(\sigma_i,\sigma_j)$ for every t' in a neighborhood of t.

We now discuss a bit more formally the extent to which player i should be willing to experiment with beliefs below p_*. Assume that $\xi_i(t) < p_*$ for some $t \in \mathbf{R}_+$, and assume first that the correlation ρ is positive. We claim that player i would rather drop shortly before t, rather than wait until t. Let us indeed place ourselves at a time $t' < t$ very close to t, and assume that player i is considering whether to experiment until t (but not until $t+$). Since $\xi_i(t) < p_*$, if player i were not taking player j's actions into account, waiting until t would yield an expected continuation payoff *below* s. More precisely, the expected *loss* would be of the order of $t - t'$.[17] Observing player j's decisions in the time interval $[t', t]$ compensates to some extent for these losses. However, the decisions of player j between t' and t make a difference to the behavior of player i only if player j happens to drop out in the time interval $[t', t)$. Given that $\rho > 0$, player i then drops out immediately following player j, rather than waits until t. By dropping out earlier, player i saves the cost of waiting, which is of the order of $t - t'$. However, the probability of that event – the probability that player j drops out between t' and t –, decreases to zero as $t' \to t$. Thus, for t' close enough to t, the marginal value of observing player j's actions is arbitrary small compared to $t - t'$ and therefore, small compared to the expected loss.

[15] conditional on $\tau_1 \geq t$

[16] and conditional on player j being active

[17] The expected loss also depend on the difference $p_* - \xi_i(t)$. This difference is here kept fixed.

We stress that this conclusion does not carry over to the case where the correlation ρ is negative. Assume indeed now that $\rho < 0$. If player j drops out at some time $x \in [t', t)$, the continuation payoff of player i is then equal to $W(\phi(x))$. If $\phi(t) > p_*$, the difference $W(\phi(x)) - s$ is bounded away from zero (for t' close to t). Thus, if the probability that player j drops out in the interval $[t', t)$ is itself of the order of $t - t'$, the marginal expected gain in observing player j's decisions is of the order of $t - t'$, and may well justify experimenting until t.

Nevertheless, there are still two cases in which the above conclusion remains valid. Firstly, if $\phi(t) \leq p_*$, and since $W \circ \phi$ is a C^1 function, the difference $W(\phi(x)) - s$ is at most of the order of $t - t'$. Then, exactly as in the case where $\rho > 0$, the expected marginal gain of observing player j is arbitrarily small compared to $t - t'$. Secondly, if the strategy σ_j is such that player j cannot possibly stop in $[t', t)$, then the marginal value of observing player j is trivially zero.

This discussion is summarized in the two observations below.

O3 Assume that $\xi_i(t) < p_*$ and $\phi(t) \leq p_*$. Then $\gamma_A^i(t, \sigma_j) < \gamma_A^i(\sigma_i, \sigma_j)$.

O4 Assume that $\xi_i(t) < p_*$ and that $t \notin S_j$. Then $\gamma_A^i(t, \sigma_j) < \gamma_A^i(\sigma_i, \sigma_j)$.

The previous observations **O1** through **O4** have useful immediate consequences, which are valid irrespective of the sign of the correlation.

Claim 5 *If $t \in S_i \setminus S_j$, then $\xi_i(t) = p_*$.*

That is, if player i finds it optimal to drop out at $t \notin S_j$, the belief of player i at t is equal to p_*. This claim follows immediately from **O2** and **O4**, noting that $\xi_i(\cdot)$ and $\gamma_A^i(\cdot, \sigma_j)$ are continuous at t since $t \notin S_j$.

Claim 6 *Assume that $\sigma_j(\{t\}) > 0$. Then $\sigma_i([t - \varepsilon, t]) = 0$ for (every) $\varepsilon > 0$ small enough.*

That is, if the strategy of player j has an atom at t, player i either stops well before t, or after t. The proof of this claim combines the different observations.

If $\gamma_A^i(t+, \sigma_j) > \gamma_A^i(t, \sigma_j)$, the conclusion follows from **O1**. Assume now that $\gamma_A^i(t+, \sigma_j) = \gamma_A^i(t, \sigma_j)$. In that case, one has $\phi(t) \leq p_*$ using (4), and $\xi_i(t) \neq \xi_i(t+)$. Note that the belief held by player i at time $t+$ is either equal to $\phi(t)$ if player j drops out, or to $\xi_i(t+)$, if player j does not drop out. Since the 'prior' belief $\xi_i(t)$ at time t is equal to the expectation of the belief held at time $t+$ (this is the martingale property of beliefs), it must be that $\xi_i(t)$ belongs to the interval $[\phi(t), \xi_i(t+)]$. This implies that either $\xi_i(t) < p_*$, or $\xi_i(t+) > p_*$. The conclusion will follow using **O3** in the former case, and from **O2** in the latter.

3.2.2 The case of good news

We here assume that the correlation between the two arms is negative. The observations made in the previous section allow us to drastically reduce the range of potential equilibria, as we now show.

Claim 7 *Assume $\rho < 0$. Then $S_i = S_j$ and the common set is a compact interval of the form $[T_p, \hat{T}]$, where $\hat{T}$ is such that $\phi(\hat{T}) > p_*$. Moreover, none of the two strategies σ_i and σ_j can possibly have an atom in $(T_p, \hat{T}]$, and at most one of the two has an atom at T_p.*

The final assertion on the atoms follows immediately from the other assertions, using Claim 6.

Since $\rho < 0$, the belief $\xi_i(t)$ is decreasing with time. We first argue, by means of contradiction, that $S_i = S_j$. Assume instead that $t_0 \in S_i \backslash S_j$, for some $t_0 \in \mathbf{R}_+$. By **O4**, one thus has $\xi_i(t_0) = p_*$, and by **O2**, one has $t_0 = \min S_i$. That is, player i cannot possibly stop prior to t_0, so that the belief held by player is $\xi_j(t) = p(t)$ for every $t < t_0$. Since the correlation between the two risky arms is negative, one has $p(t) \leq \xi_i(t)$ for every $t < t_0$ and therefore, $\xi_j(t) > p_*$ for every such t. Hence, $t_0 < \min S_j$ by **O2**. This implies in turn that $p(t_0) = \xi_i(t_0)(= p_*)$.

Next, set $t_1 := \min\{t > t_0 : t \in S_i \cup S_j\}$. It is not difficult to combine the previous remarks to show that $t_1 > t_0$ and $t_1 \in S_i \cap S_j$ so that $t_1 = \min S_j$. Arguing as above, one then has $\xi_j(t) < \xi_i(t) = p(t) < p_*$ for each $t \in (t_0, t_1]$. By Claim 6, one at least of the two strategies σ_i and σ_j – say, σ_2 –, assigns probability zero to t_1. This implies that $\gamma^1(t_1, \sigma_2)$ is equal to the equilibrium payoff $\gamma^1(\sigma_1, \sigma_2)$, in contradiction with **O4**. This implies $S_i = S_j$, as claimed. We denote by S the common set.

Next, using both Claim 6, **O2** and **O3**, it is easy to check that $T_p = \min S$. If S failed to be an interval, there would exist $t_0, t_1 \in S$, with $T_p \leq t_0 < t_1$ and $(t_0, t_1) \cap S = \emptyset$. We now essentially repeat the previous paragraph. By Claim 6, there is a player, say j, such that $\sigma_j(\{t_1\}) = 0$, which implies $\gamma^i_A(t_1, \sigma_j) = \gamma^i_A(\sigma_i, \sigma_j)$, in contradiction with **O4**.

In order to complete the discussion of Theorem 4, we argue informally that there is at most one equilibrium. (Existence is proven in the Appendix.) Let as above an equilibrium (σ_1, σ_2) be given. Using Claim 7, we denote by $[T_p, \hat{T}]$ the common support of the two distributions σ_1 and σ_2. Wlog, we may further assume that σ_2 is a non-atomic distribution. This implies that the indifference condition $\gamma^1_A(t, \sigma_2) = \gamma^1_A(\sigma_1, \sigma_2)$ holds for each $t \in [T_p, \hat{T}]$. We will first show that this indifference condition uniquely pins down both the endpoint $\hat{T}$ and the distribution σ_2. We next claim that σ_1 has to be equal to σ_2, thereby establishing uniqueness.

We describe the distribution σ_2 by its cdf, $F_2(t) = \sigma_2([0, t]) = \sigma_2([T_p, t])$ $(t \in [T_p, \hat{T}])$. Note that $F_2(t)$ is also equal to $\mathbf{P}(\theta_2 \leq t \mid R_2 = B)$, and that the conditional probability

$\tilde{F}_2(t) := \mathbf{P}(\theta_2 \leq t \mid R_2 = G)$ is related to F_2 through the identity

$$\begin{aligned} 1 - \tilde{F}_2(t) &= \mathbf{P}(\theta_2 \leq t, \leq \tau_2 \mid R_2 = G) + \mathbf{P}(\theta_2 \leq \tau_2 \leq\mid R_2 = G) \\ &= e^{-\lambda t}(1 - F_2(t)) + \int_0^t \lambda e^{-\lambda t}(1 - F_2(z))dz. \end{aligned}$$

We place ourselves at some time $t \in [T_p, \hat{T}]$, and will write an expansion of the equality $\gamma_A^1(t + dt, \sigma_2) = \gamma_A^1(t, \sigma_2)$, as $dt \to 0$. If either $\tau_1 < t$ or $\theta_2 < t$, the outcomes induced by the two profiles (t, σ_2) and $(t + dt, \sigma_2)$ are the same. We henceforth condition on the event $\min(\tau_1, \theta_2) \geq t$.

Consider the three events $\mathcal{A} := \{\tau_1 < t + dt\}$, $\mathcal{B} := \{\tau_1 \geq t + dt, \theta_2 \in [t, t + dt)\}$, and $\mathcal{C} := \{\tau_1 \geq t + dt, \theta_2 \geq t + dt\}$. (Conditional on $\min(\tau_1, \theta_2) \geq t$), the probability of $\mathcal{A}$ is of the order of $\lambda \xi_1(t) dt$, while the probability of $\mathcal{B}$ is $\mathbf{P}(\theta_2 \in [t, t + dt) \mid \min(\tau_1, \theta_2) \geq t)$. The continuation payoff induced by $(t + dt, \sigma_2)$ as of time t, is approximately equal γ on the event $\mathcal{A}$, equal to $W(\phi(t))$ on the event $\mathcal{B}$, and is exactly equal to se^{-rdt} on the event $\mathcal{C}$, while the continuation payoff induced by (t, σ_2) is equal to s in all three cases.

Thus, the equality $\gamma_A^1(t + dt, \sigma_2) = \gamma_A^1(t, \sigma_2)$ writes

$$(\gamma - s)\lambda \xi_1(t) + (W(\phi(t)) - s)\frac{\mathbf{P}(\theta_2 \in [t, t + dt) \mid \min(\tau_1, \theta_2) \geq t)}{dt} = rs.$$

This 'shows' that the limit $\mu_1(t) := \lim_{dt \to 0} \frac{\mathbf{P}(\theta_2 \in [t, t+dt) \mid \min(\tau_1, \theta_2) \geq t)}{dt}$ is well-defined and solves

$$(\gamma - s)\lambda \xi_1(t) + (W(\phi(t)) - s)\mu_1(t) = rs. \tag{5}$$

The value of the limit $\mu_1(t)$ may be interpreted as a conditional hazard rate for θ_2, conditional on $\tau_1 \geq t$. Note that the probability $\mathbf{P}(\tau_1 \geq t, \theta_2 \in [t, t + dt))$ is given by

$$\left(q_{GG}e^{-\lambda t} + q_{GB}\right)(\tilde{F}_2(t + dt) - \tilde{F}_2(t)) + \left(q_{GB}e^{-\lambda t} + q_{BB}\right)(F_2(t + dt) - F_2(t)),$$

where $q_{\omega_1 \omega_2}$ stands for the prior probability that $R_1 = \omega_1$ and $R_2 = \omega_2$). A similar formula holds for $\mathbf{P}(\tau_1 \geq t, \theta_2 \geq t)$. Letting $dt \to 0$ and plugging into (5), standard algebraic manipulations then show that F_2 solves the integro-differential equation (1) on the interval $[T_p, \hat{T}]$.

Since σ_2 is concentrated on $[T_p, \hat{T}]$, one has moreover $F_2(T_p) = 0$ and $\displaystyle\int_0^{T_p} e^{-\lambda x}(1 - F(x)) = \frac{1 - e^{-\lambda T_p}}{\lambda}$. There is a unique solution, say $\hat{F}_2$, to the equation (1) on $[T_p, T_\phi]$, that satisfies these initial conditions. Therefore, it must be that $\hat{T} := \min\{t : \hat{F}_2(t) = 1\}$, and F_2 coincides with $\hat{F}_2$ on the interval $[T_p, \hat{T}]$. This shows the uniqueness of $\hat{T}$ and of F_2, as desired.

We turn to the strategy σ_1. Since σ_1 has no atom in the interval $(T_p, \hat{T}]$, one has $\gamma_A^2(t, \sigma_1) = \gamma_A^2(t', \sigma_1)$ for every $t, t' \in (T_p, \hat{T}]$, so that the hazard rate μ_2 of player 1's exit decision is related to player 2's belief ξ_2 through the equation (5), for every $t \in (T_p, \hat{T}]$. The economic intuition behind the uniqueness claim is as follows.

Assume for a moment that σ_1 has an atom located at T_p: $\sigma_1(\{T_p\} > 0$. Then, when at time T_p+ (if player 1 remains active), player 2 is *more* pessimistic than player 1: $\xi_2(T_p+) < \xi_1(T_p+)$. Since both players 1 and 2 are indifferent between dropping out immediately, and waiting a further infinitesimal amount of time, it must be that player 2 assigns *higher* chances than player 1 does, to the fact that he will receive good news in this time interval. In other words, the probability of player 1 dropping out should be higher than the probability of player 2 dropping out, to compensate for the fact that player 2 is more pessimistic. Indeed, from (5), one has $\mu_2(t) > \mu_1(t)$ as soon as $\xi_1(t) > \xi_2(t)$. Should players remain active, this will make player 2 even more pessimistic than player 1. Proceeding "inductively", one obtains more generally that $F_1(t) > F_2(t)$ for all $t \in [T_p, \hat{T}]$ and therefore, $F_1(t) = 1$ for some $t < \hat{T}$ – in contradiction with the fact that σ_1 and σ_2 have the same support.

3.2.3 The case of bad news

We now turn to the case where the correlation between the two risky arms is positive. Again, we derive direct implications of the observations **O1-4** and of Claims 5 and 6.

We first check that the main qualitative insight is valid (Proposition 2). Let (σ_1, σ_2) be an equilibrium and assume by way of contradiction that $\xi_i(t) < p_*$ for some $t \in S_i$. Since $\rho > 0$, the two beliefs $\phi(t)$ and $\xi_i(t)$ satisfy $\phi(t) \leq \xi_i(t)$. Using (4), this yields $\gamma_A^i(t+, \sigma_j) = \gamma_A^i(t, \sigma_j)$. This implies in turn that the payoff $\gamma_A^i(t, \sigma_j)$ induced by the strategy t must be equal to the equilibrium payoff $\gamma_A^i(\sigma_i, \sigma_j)$, in contradiction with **O3**. Thus, whenever player i is willing to drop out, his belief ξ_i is equal to the one-player threshold, p_*.

Uniqueness of the symmetric equilibrium

Let (σ, σ) be a symmetric equilibrium, with support S. Using Claim 6, the distribution σ is necessarily non-atomic. By the Proposition 2 just proven, one must have $\xi_i(t) = p_*$ for each $t \in S$. The only way for a player's belief to remain constant over time is when the bad news coming from one's own arm is exactly offset by the good news coming from the other player, at any point in time. This suggests that the support S of σ must be an interval, say $[T_1, T_2]$.

Note now that $\xi_i(T_1) = p(T_1)$: indeed, since player j never stops prior to T_1, the event $\theta_j \geq T_1$ is uninformative. On the other hand, player j stops for sure before T_2, unless he has hit a success: thus, $\xi_i(T_2) = \psi(T_2)$. Since $\xi_i(T_1) = \xi_i(T_2) = p_*$, T_1 and T_2 must therefore be equal to T_p and T_ψ respectively. Finally, the distribution σ is uniquely defined by the condition that $\xi_i(t) = p_*$ for each $t \in [T_p, T_\psi)$. While this establishes uniqueness, this does not prove existence of a symmetric equilibrium. The existence proof is common to the case where $\rho < 0$, and is postponed to the Appendix.

Pure equilibria

In contrast with the case where $\rho > 0$, there may be many equilibria with finite support, as stated in Proposition 3. Let first (t_1, t_2) be a tentative pure equilibrium. By Claim 6, one has $t_1 \neq t_2$. Assume wlog that $t_1 < t_2$. As above, one then has $\xi_1(t_1) = p(t_1)$ and $\xi_2(t_2) = \psi(t_2)$ and therefore, $t_1 = T_p$ and $t_2 = T_\psi$. We now argue that (T_p, T_ψ) is indeed a pure equilibrium.

When facing T_p, the payoff function $t \mapsto \gamma_A^2(t, \sigma_1)$ of player 2 is increasing over $[0, T_\psi]$ since $\xi_2(t+) > p_*$ for every $t < T_\psi$, and decreasing beyond T_ψ, with a discontinuity at T_p. Therefore, T_ψ is the unique best reply of player 2 to T_p.

When facing T_ψ, and for $t \leq T_\psi$, the payoff of player 1 is the same as in the decision problem $\mathcal{P}$, since player 1 does not deduce any information from the behavior of player 2 up to time t. This payoff is therefore increasing over $[0, T_p]$, and decreasing on $[T_p, T_\psi]$. Should player 1 choose to wait beyond T_ψ, and to learn the decision of player 2 at time T_ψ, his belief would then jump either to $\phi(T_\psi)$ or to $\psi(T_\psi)$, depending on that decision. But since $\phi(T_\psi) < \psi(T_\psi) = p_*$, player 1 will then find it optimal to drop out anyway. Hence T_p is the unique best reply to the strategy T_ψ.

Aymmetric equilibria

The proof that equilibria may exist with a finite support of arbitrary size is a bit delicate and we briefly discuss the main ideas. Let a tentative equilibrium (σ_1, σ_2) be given, where σ_i assigns positive probability to exactly k dates. By Claim 6, the two equilibrium supports S_1 and S_2 are disjoint. Since $\xi_i(t_i) = p_*$, it must be that the two sets S_1 and S_2 are *interlaced*: any two consecutive dates in S_i are separated by exactly one date in S_j. The equilibrium logic is as follows. If player i does not drop out at some $t_i \in S_i$, this is good news for player j, whose belief jumps above p_*, and then declines continuously until it reaches p_* at the next date t_j in S_j. Meanwhile, player i's belief declines below p_* until date t_j. If player j does not drop out at that date, player i' s belief jumps above p_*, and then declines continuously until it reaches p_* at the next date in S_i. And so on.

We illustrate the proof in the specific case where $k = 2$, and $S_i = \{\underline{t}_i, \bar{t}_i\}$. Assuming wlog $\underline{t}_1 < \underline{t}_2$, the same arguments as above imply $\underline{t}_1 = T_p$ and $\bar{t}_2 = T_\psi$. We are thus left with four unknowns – $\underline{t}_2$ and $\bar{t}_1$, and the weight assigned by σ_i to $\underline{t}_i-$, and four equilibrium conditions: the belief conditions $\xi_1(\bar{t}_1) = \xi_2(\underline{t}_2) = p_*$ and the indifference condition for player i at time $\underline{t}_i$. Existence is proven using a fixed point argument in the convex and compact set of pairs $(\underline{t}_2, \bar{t}_1)$ such that $T_p \leq \underline{t}_2 \leq \bar{t}_1 \leq T_\psi$. In a nutshell, the proof goes as follows. Given $\bar{t}_1 \in [T_p, T_\psi]$, we prove that there is a unique $\underline{t}_2 \in [T_p, \bar{t}_1]$ and a unique weight on $\underline{t}_2$ such that (i) the belief held by player 1 at $\bar{t}_1$ is equal to t_* and (ii) the payoffs induced by the two strategies T_p and $\bar{t}_1$ are the same. Moreover, $\underline{t}_2$ is continuous w.r.t. $\bar{t}_1$ and $\underline{t}_2 \in (T_p, \bar{t}_1)$ whenever $\bar{t}_1 > T_p$. Using a similar argument, we prove that for given $\underline{t}_2$, there is a unique $\bar{t}_1 \in [\underline{t}_2, T_\psi]$ and a unique weight on $\underline{t}_1 = T_p$ such that (i) the belief of player 2 at $\underline{t}_2$ is p_* and (ii) the payoffs induced by the two strategies $\underline{t}_2$ and T_ψ are the same.

References

[1] P. Aghion, P. Bolton, C. Harris, and B. Jullien. Optimal learning by experimentation. *The Review of Economic Studies*, **58**(4): 621–654, 1991.

[2] D. Bergemann and U. Hege. The financing of innovation: Learning and stopping. *The RAND Journal of Economics*, **36**(4): 719–752, 2005.

[3] D. Bergemann and J. Valimaki. Bandit Problems. in *The New Palgrave Dictionary of Economics*, 2nd edition, editors S. Durlauf and L. Blume. Basingstoke and New York: Palgrave Macmillan Ltd, 2008.

[4] P. Bolton and C. Harris. Strategic experimentation. *Econometrica*, **67**(2): 349–374, 1999.

[5] G. Keller, S. Rady. Strategic experimentation with Poisson bandits. *Theoretical Economics*, **5**: 275–311, 2010.

[6] G. Keller, S. Rady, and M. Cripps. Strategic experimentation with exponential bandits. *Econometrica*, **73**(1): 39–68, 2005.

[7] N. Klein and S. Rady. Negatively correlated bandits. Working paper, University of Munich, 2010.

[8] A. McLennan. Price Diversity and Incomplete Learning in the Long Run. *Journal of Economic Dy- namics and Control*, **7**: 331-347, 1984.

[9] D. A. Malueg and S. O. Tsutsui. Dynamic R&D competition with learning. *The RAND Journal of Economics*, **28**(4): 751–772, 1997.

[10] G. Moscarini and F. Squintani. Competitive experimentation with private information: The survivor's curse. *Journal of Economic Theory*, **145**: 639–660, 2010.

[11] P. Murto and J. Välimäki. Learning and information aggregation in an exit game. Mimeo, 2009.

[12] E.L. Presman. Poisson version of the two-armed bandit problem with discounting. *Theory of Probability and its Applications*, **35**, 307–317, 1990.

[13] E.L. Presman and I.N. Sonin. *Sequential control with incomplete information: the Bayesian Approach to multi-armed bandit problems.* London: Academic Press, 1990.

[14] D. Rosenberg, E. Solan, and N. Vieille. Social learning in one-arm bandit problems. *Econometrica*, **75**(6): 1591–1611, 2007.

Appendix

We here provide all the formal details. We start with the analysis of the decision problem $\tilde{\mathcal{P}}$ In Section A, and next complete the analysis of the payoffs scenario. The much longer section B contains the analysis of the actions scenario.

A The decision problem $\tilde{\mathcal{P}}$

We adopt the convention that the decision maker in $\tilde{\mathcal{P}}$ is player 2. We recall that the expected payoff induced in $\tilde{\mathcal{P}}$ by the pure strategy t is denoted by $\tilde{\pi}(t)$, and is given by

$$\tilde{\pi}(t) = \mathbf{E}\left[e^{-r\tau_1}W(\psi(\tau_1))1_{\tau_1\leq\min(\tau_2,t)}\right] + \gamma\mathbf{E}\left[e^{-r\tau_2}1_{\tau_2<\min(\tau_1,t)}\right] + se^{-rt}\mathbf{P}(\min(\tau_1,\tau_2)\geq t), \text{ for } t\in\mathbf{R}_+.$$

Thus, for every two dates $t < t'$, one has

$$\begin{aligned}\tilde{\pi}(t') - \tilde{\pi}(t) &= \mathbf{E}\left[e^{-r\tau_1}W(\psi(\tau_1))1_{t\leq\tau_1<\min(\tau_2,t')}\right] + \gamma\mathbf{E}\left[e^{-r\tau_2}1_{t\leq\tau_2<\min(\tau_1,t')}\right] &(6)\\ &\quad +s\left(e^{-rt'}\mathbf{P}(\min(\tau_1,\tau_2)\geq t') - e^{-rt}\mathbf{P}(\min(\tau_1,\tau_2)\geq t)\right). &(7)\end{aligned}$$

We will prove that there is a unique optimal strategy, T_{**}, which is obtained as the unique solution to the equation

$$\lambda\phi(T_{**})\left(W(\psi(T_{**}) + \gamma - 2s\right) = rs.$$

In addition, $T_{**}\geq T_\phi$, and $T_{**} > T_\phi$ if and only if $\rho > 0$.

The proof is organized in several steps.

Claim 8 *There exists an optimal strategy.*

Proof of the claim. By dominated convergence, $\tilde{\pi}$ is continuous over $\mathbf{R}_+$, and has a limit when $t\to+\infty$, which we denote $\tilde{\pi}(+\infty)$. Letting $t'\to+\infty$ in (6), one gets

$$\tilde{\pi}(+\infty) - \tilde{\pi}(t) = \mathbf{E}\left[e^{-r\tau_1}W(\psi(\tau_1))1_{t\leq\tau_1<\tau_2}\right] + \gamma\mathbf{E}\left[e^{-r\tau_2}1_{t\leq\tau_2<\tau_1}\right] - se^{-rt}\mathbf{P}(\min(\tau_1,\tau_2)\geq t). \quad (8)$$

The first expectation in the right-hand side of (8) is at most $\gamma e^{-rt}\mathbf{P}(t\leq\tau_1<+\infty)$, while the second one is at most $\gamma e^{-rt}\mathbf{P}(t\leq\tau_2<+\infty)$. The probabilities $\mathbf{P}(t\leq\tau_1<+\infty) = \mathbf{P}(t\leq\tau_2<+\infty)$ converge to zero as $t\to+\infty$, while $\mathbf{P}(\min(\tau_1,\tau_2)\geq t)$ has a positive limit. Thus, $\tilde{\pi}(+\infty) < \tilde{\pi}(t)$ for every large t. It follows that $\tilde{\pi}$ admits (at least) one maximum. ■

Claim 9 *The function $\tilde{\pi}$ is of class C^1 and its derivative is given by*

$$\tilde{\pi}'(t) = e^{-rt}\mathbf{P}(\min(\tau_1, \tau_2) \geq t) \times \{\lambda\phi(t)\left(W(\psi(t)) + \gamma - 2s\right) - rs\}, t \in \mathbf{R}_+. \tag{9}$$

Proof. Fix $t_0 \in \mathbf{R}_+$. We will prove the claim relative to the right-derivative. Similar arguments apply for the existence of a left derivative.

Observe first that, for $t > t_0$,

$$\begin{aligned} \mathbf{P}(\tau_1, \tau_2 \in [t_0, t]) &\leq \mathbf{P}(\tau_1, \tau_2 \in [t_0, t] \mid R_1 = R_2 = G) \\ &= e^{-2\lambda t}\left(1 - e^{-\lambda(t-t_0)}\right)^2 \leq \lambda^2(t - t_0)^2. \end{aligned}$$

Therefore,

$$\lim_{t \searrow t_0} \frac{\mathbf{P}(\tau_1, \tau_2 \in [t_0, t])}{t - t_0} = 0. \tag{10}$$

Thus, when taking the limit of $\dfrac{\tilde{\pi}(t) - \tilde{\pi}(t_0)}{t - t_0}$ (see (6)), the indicators $1_{t_0 \leq \tau_1 < \min(\tau_2, t)}$ and $1_{t_0 \leq \tau_2 < \min(\tau_1, t)}$ which appear in the first two expectations can be replaced with $1_{t_0 \leq \tau_1 < t}$ and $1_{t_0 \leq \tau_2 < t}$ respectively.

Next, consider each of the (simplified) two expectations in turn (divided by $\mathbf{P}(\min(\tau_1, \tau_2) \geq t_0)$). Observe first that

$$\frac{1}{t - t_0}\mathbf{E}\left[e^{-r(\tau_1 - t_0)}W(\psi(\tau_1))1_{\tau_1 \in [t_0, t)} \mid \min(\tau_1, \tau_2 \geq t_0)\right] = \phi(t_0)\frac{1}{t - t_0}\int_{t_0}^{t} e^{-r(x - t_0)}W(\psi(x))\lambda e^{-\lambda(x - t_0)}dx,$$

which converges to $\lambda\phi(t_0)W(\psi(t_0))$ when $t \searrow t_0$, since $W(\cdot)$ and $\psi(\cdot)$ are continuous.

Observe next that $\dfrac{1}{t - t_0}\mathbf{E}\left[e^{-r(\tau_2 - t_0}1_{\tau_2 \in [t_0, t)} \mid \min(\tau_1, \tau_2 \geq t_0\right]$ converges to $\lambda\phi(t_0)$, for similar reasons.

We finally deal with the last term in the expression of $\tilde{\pi}(t) - \tilde{\pi}(t_0)$ in (6). One has

$$\begin{aligned} & \frac{1}{\mathbf{P}(\min(\tau_1, \tau_2) \geq t_0)} \times \frac{1}{t - t_0}\left(e^{-r(t - t_0)}\mathbf{P}\left(\min(\tau_1, \tau_2) \geq t\right) - \mathbf{P}(\min(\tau_1, \tau_2) \geq t_0)\right) \\ = \ & \frac{1}{t - t_0}\left(e^{-r(t - t_0)}\mathbf{P}(\min(\tau_1, \tau_2) \geq t \mid \min(\tau_1, \tau_2) \geq t_0) - 1\right) \\ = \ & \frac{1}{t - t_0}\left((e^{-r(t - t_0)} - 1)\mathbf{P}\left(\min(\tau_1, \tau_2) \geq t \mid \min(\tau_1, \tau_2) \geq t_0\right) - \mathbf{P}\left(\min(\tau_1, \tau_2) < t \mid \min(\tau_1, \tau_2) \geq t_0\right)\right) \end{aligned}$$

The first term converges to $-r$. Using (10), the second term has the same limit, namely $-2\lambda\phi(t_0)$, as the expression

$$-\left(\mathbf{P}(\tau_1 \in [t_0, t) \mid \min(\tau_1, \tau_2) \geq t_0) + \mathbf{P}(\tau_2 \in [t_0, t) \mid \min(\tau_1, \tau_2) \geq t_0)\right).$$

The result follows when adding these different limits. ■

We now conclude the proof of Proposition 1. Let $T_{**} \in \mathbf{R}_+$ be any point at which $\tilde{\pi}$ reaches its maximum. One has $\tilde{\pi}'(T_{**}) = 0$, that is

$$\lambda\phi(T_{**})W(\psi(T_{**})) + \lambda\phi(T_{**})\gamma - s(r + 2\lambda\phi(T_{**})) = 0$$

or equivalently,

$$\lambda\phi(T_{**})\left(W(\psi(T_{**}) + \gamma - 2s\right) = rs. \tag{11}$$

Since ϕ is decreasing, positive and continuous, and since $W(\psi(\cdot))$ is non-increasing, Equation (11) has a unique solution.

Recall now that T_ϕ solves

$$\lambda\phi(T_\phi)(\gamma - s) = rs.$$

Since $W(\psi(T_{**})) \geq s$, this implies $\phi(T_{**}) \leq \phi(T_\phi)$, hence $T_{**} \geq T_\phi$.

If $\rho < 0$, one has $\psi(T_{**}) \leq \phi(T_{**}) \leq \phi(T_\phi) = p_*$, hence $W(\psi(T_{**})) = s$, so that Equation (3) reduces in that case to $\lambda\phi(T_{**})(\gamma - s) = rs$, and $T_{**} = T_\phi$.

If instead $\rho > 0$, then $\psi(T_*) > \phi(T_\phi)$, hence $W(\psi(T_\phi)) > s$, so that T_ϕ is not a solution to (11). In that case, one therefore has $T_{**} > T_\phi$.

B The actions scenario

In this section, we provide the proofs of all the results relative to the actions scenario. The section is organized as follows. We first introduce in Section B.1 some notation and prove a few properties, which are valid irrespective of the sign of the correlation. Next, we prove that the marginal value of the informational externality is zero when $\rho > 0$ (Proposition 1 in the paper), and characterize the unique symmetric equilibrium in that case (Theorem 1 in the paper). We then prove the existence and the uniqueness of the equilibrium when $\rho < 0$ (Theorem 4 in the paper), and provide a characterization of the equilibrium. Finally, we come back to the case of a positive correlation, and prove the multiplicity of asymmetric equilibria (Proposition 2 in the paper).

B.1 Technical Preliminaries

We let a strategy σ_j of player j be fixed throughout this section. Recall that the belief function ξ_i is defined by $\xi_i(t) = \mathbf{P}(R^i = G \mid \tau_i \geq t, \theta_j \geq t)$ $(t \in \mathbf{R}_+)$, which is computed under the assumption that player j is using the strategy σ_j. It is readily checked that $\xi_i(\cdot)$ is left-continuous over $\mathbf{R}_+$, and has a right-limit at each $t \in \mathbf{R}_+$, which is given by $\xi_i(t+) := \mathbf{P}(R^i =$

$G \mid \tau_i > t, \theta_j > t)$. In addition, ξ_i is continuous at $t \in \mathbf{R}_+$ if and only if $\mathbf{P}(\theta_j = t) = 0$ that is, if and only if $\sigma_j(\{t\}) = 0$.

B.1.1 Expected Payoffs

This section serves as a reminder for some notation, and provides a few elementary properties of the payoff function γ_A.

When facing the strategy σ_j, and when using the strategy t, the continuation payoff of player i at time $\min(t, \tau_i, \theta_j)$ is equal to (i) s if $\theta_j \geq t$ and $\tau_i \geq t$, (ii) γ if $\tau_i \leq \theta_j$ and $\tau_i < t$, and (iii) $W(\phi(\theta_j))$ if $\theta_j < \min(t, \tau_i)$.

Since τ_i has a density conditional on $R_i = G$, one has $\mathbf{P}(\tau_i = \theta_j < +\infty) = 0$, and player i's expected payoff under (t, σ_j) is thus given by

$$\gamma_A^i(t, \sigma_j) = se^{-rt}\mathbf{P}(\theta_j \geq t, \tau_i \geq t) + \gamma\mathbf{E}\left[e^{-r\tau_i}1_{\tau_i < \min(\theta_j, t)}\right] + \mathbf{E}\left[e^{-r\theta_j}W(\phi(\theta_j))1_{\theta_j < \min(\tau_i, t)}\right].$$

In particular, for $t_0 < t$, one has

$$\gamma_A^i(t, \sigma_j) - \gamma_A^i(t_0, \sigma_j) = se^{-rt}\mathbf{P}(\theta_j \geq t, \tau_i \geq t) - se^{-rt_0}\mathbf{P}(\theta_j \geq t_0, \tau_i \geq t_0) \tag{12}$$
$$+\gamma\mathbf{E}\left[e^{-r\tau_i}1_{t_0 \leq \tau_i < \min(\theta_j, t)}\right] + \mathbf{E}\left[e^{-r\theta_j}W(\phi(\theta_j))1_{t_0 \leq \theta_j < \min(\tau_i, t)}\right]. \tag{13}$$

By dominated convergence, the map $t \mapsto \gamma_A^i(t, \sigma_j)$ is left-continuous over $\mathbf{R}_+$, and admits a right-limit at each $t \in \mathbf{R}_+$, which we denote $\gamma_A^i(t+, \sigma_j)$, and which is given by

$$\gamma_A^i(t+, \sigma_j) - \gamma_A^i(t, \sigma_j) = e^{-rt}\mathbf{P}(\theta_j = t, \tau_i \geq t) \times (W(\phi(t)) - s). \tag{14}$$

Since $W(p) \geq s$ for each $p \in [0, 1]$, this implies that $\gamma_A^i(t+, \sigma_j) \geq \gamma_A^i(t, \sigma_j)$, and that equality holds when either $\sigma_j(\{t\}) = 0$, or $\phi(t) \leq p_*$.

If σ_i is a best reply of player i to σ_j, then $\gamma_A^i(t, \sigma_j) = \gamma_A^i(\sigma_i, \sigma_j)$ for σ_i-a.e. $t \in [0, +\infty]$. In particular, for a given $t \in \mathbf{R}_+$, one has $\gamma_A^i(t, \sigma_j) = \gamma_A^i(\sigma_i, \sigma_j)$ whenever $\sigma_j(\{t\}) = 0$, and t belongs to the support of σ_i. We will also rely on the following fact.

Fact 0. For every t in the support of σ_i, one has $\gamma_A^i(t+, \sigma_j) = \gamma_A^i(\sigma_i, \sigma_j)$.

Indeed, note that any such t is a limit of a sequence (t_n) such that[18] $\gamma_A^i(t_n, \sigma_j) = \gamma_A^i(\sigma_i, \sigma_j)$. Note next that $\limsup_{n \to +\infty} \gamma_A^i(t_n, \sigma_j) \leq \gamma_A^i(t+, \sigma_j)$ for any sequence $t_n \to t$. This yields **Fact 0**.

It will be convenient to study and characterize equilibria by means of the differential properties

[18] Otherwise indeed, one would have $\gamma_A^i(x, \sigma_j) < \gamma_A^i(\sigma_i, \sigma_j)$ for all x in some neighborhood of t, and player i would be better off assigning probability zero to that neighborhood.

of $t \mapsto \gamma_A^i(t, \sigma^j)$. We here gather a few simple observations, to which we will repeatedly refer later. All these observations are valid, irrespective of the strategy σ_j.

We let $x > 0$ be given, and will let $y > x$ vary.[19]

Since τ_i follows an exponential distribution if $R_i = G$, one has:

Fact 1. $\lim_{y \searrow x} \frac{1}{y-x} \mathbf{P}(\tau_i < y \mid \min(\theta_j, \tau_i) \geq x) = \lambda \xi_i(x)$.

Since $\mathbf{P}(\theta_j \in (x, y))$ converges to zero as $y \to x$, and since θ_j and τ_i are conditionally independent given the types (R^i, R^j), **Fact 1** implies that

$$\lim_{y \searrow x} \frac{1}{y-x} \mathbf{P}(x < \theta_j, \tau_i \leq y) = 0. \tag{15}$$

Next, observe that $e^{-r(y-x)} \mathbf{P}(\min(\theta_j, \tau_i) \geq y \mid \min(\theta_j, \tau_i) \geq x) - 1$ is equal to

$$(e^{-r(y-x)} - 1) \mathbf{P}(\min(\theta_j, \tau_i) \geq y \mid \min(\theta_j, \tau_i) \geq x) - \mathbf{P}(\min(\theta_j, \tau_i) < y \mid \min(\theta_j, \tau_i) \geq x).$$

Combining (15) with **Fact 1**, this easily yields **Fact 2** below.

Fact 2. The sum of $\frac{1}{y-x} \left(e^{-r(y-x)} \mathbf{P}(\min(\theta_j, \tau_i) \geq y \mid \min(\theta_j, \tau_j) \geq x) - 1 \right)$ and of $\frac{1}{y-x} \mathbf{P}(\theta_j < y \mid \min(\theta_j, \tau_j) \geq x)$ converges to $r + \lambda \xi^i(x)$ as $y \searrow x$.

Let $\beta(\cdot)$ be a C^1 function of time. For $y > x$ close to x, the difference $\sup_{t \in [x,y]} |\beta(t) - \beta(x)|$ is at most of the order of $(y - x)\beta'(x)$. Since $\lim_{y \to x} \mathbf{P}(\theta_j \in (x, y)) = 0$, and together with the previous fact, this implies **Fact 3** below.

Fact 3. Let β be a C^1 function. The difference between $\frac{1}{y-x} \mathbf{E}[\beta(\theta_j) 1_{\theta_j < \min(y, \tau_i)} \mid \min(\tau_i, \theta_j) \geq x]$ and $\frac{1}{y-x} \beta(x) \mathbf{P}(\theta_j < y \mid \min(\tau_i, \theta_j) \geq x)$ converges to zero as $y \to x$.

For similar reasons, one has **Fact 3b** below.

Fact 3b. Let β be a C^1 function. The difference between $\frac{1}{y-x} \mathbf{E}[\beta(\tau_i) 1_{\tau_i < \min(y, \theta_j)} \mid \min(\tau_i, \theta_j) \geq x]$ and $\frac{1}{y-x} \mathbf{E}[\beta(\tau_i) 1_{\tau_i < y} \mid \min(\tau_i, \theta_j) \geq x)$ converges to zero as $y \searrow x$. In addition, $\lim_{y \searrow x} \frac{1}{y-x} \mathbf{E}[\beta(\tau_i) 1_{\tau_i < \min(y, \theta_j)} \mid \min(\tau_i, \theta_j) \geq x] = \lambda \xi_i(x) \beta(x)$.

[19] We assume without further notice that $\mathbf{P}(\theta_j > x) > 0$.

B.1.2 The actions game and the decision problem $\mathcal{P}$

We here discuss the relation between optimal policies in $\mathcal{P}$, and monotonicity properties of the map $t \mapsto \gamma_A^i(t,\sigma_j)$. Let first I be an interval that is assigned probability zero by σ_j. Then, while in the time interval I, player i is 'alone' and the comparison between two policies $\underline{t}$ and t in I is driven by the map W. Formally, let $t_0, \underline{t}, t \in I$ be such that $t_0 < \underline{t} < t$. By (12), one has

$$\frac{\gamma_A^i(t,\sigma_j) - \gamma_A^i(\underline{t},\sigma_j)}{e^{-r\underline{t}}\mathbf{P}(\tau_i \geq \underline{t}, \theta_j \geq t_0)} = se^{-r(t-\underline{t})}\mathbf{P}\left(\tau_i \geq t \mid \tau_i \geq \underline{t}, \theta_j \geq t_0\right) + \gamma \mathbf{E}\left[e^{-r(\tau_i - \underline{t})}1_{\tau_i < t} \mid \tau_i \geq \underline{t}, \theta_j \geq t_0\right] - s. \tag{16}$$

The right-hand side of equation (16) is equal to the payoff induced in the decision problem $\mathcal{P}$ by the policy $t - \underline{t}$, when starting with a belief equal to $\xi_i(\underline{t})$. This immediately implies the following fact.

Fact 4. Under the above assumptions, one has $\gamma_A^i(t,\sigma_j) < \gamma_A^i(\underline{t},\sigma_j)$ if $\xi_i(\underline{t}) < p_*$, and $\gamma_A^i(t,\sigma_j) > \gamma_A^i(\underline{t},\sigma_j)$ if $\xi_i(t) > p_*$.[20]

Our next observation has a straightforward intuition, but the formal proof requires a few details.

Fact 5. Assume that $\xi_i(t_0+) > p_*$. Then $\gamma_A^i(t,\sigma_j) > \gamma_A^i(t_0,\sigma_j)$ for all $t > t_0$ close enough to t_0.

The intuition is as follows. We place ourselves at time t_0, in the event where $\tau_i \geq t_0$ and $\theta_j \geq t_0$. Player i may choose to exit (strategy t_0), and thereby to get a continuation payoff equal to s. Or he may choose to wait for player j's decision, and then play optimally. Since $\xi_i(t_0+) > p_*$, the latter choice yields a higher payoff.

The formal proof goes as follows. Since $\xi_i(t_0+) > p_*$, the derivative of the map $W(\cdot)$, evaluated at $\xi_i(t_0+)$, is positive. We will prove that

$$\liminf_{t \searrow t_0} \frac{\gamma_A^i(t,\sigma_j) - \gamma_A^i(t_0,\sigma_j)}{t - t_0} \geq e^{-rt_0}W'(\xi_i(t_0+))\mathbf{P}(\tau_i \geq t_0, \theta_j > t_0),$$

which yields the result.

Since $W(\phi(x)) \geq s$ for every $x \geq t_0$, one has using (12),

$$\begin{aligned}\gamma_A^i(t,\sigma_j) - \gamma_A^i(t_0,\sigma_j) \geq\ & se^{-rt}\mathbf{P}\left(\tau_i \geq t, \theta_j \geq t\right) - se^{-rt_0}\mathbf{P}\left(\tau_i \geq t_0, \theta_j \geq t_0\right) \\ & + s\mathbf{E}\left[e^{-r\theta_j}1_{t_0 \leq \theta_j < \min(\tau_i, t)}\right] + \gamma \mathbf{E}\left[e^{-r\tau_i}1_{t_0 \leq \tau_i < \min(\theta_j, t)}\right].\end{aligned}$$

[20] Note that the condition is set on $\xi_i(\underline{t})$ in the first case, and on $\xi_i(t)$ in the second case.

When moving the event $\theta_j = t_0$ from the third to the second term, and using the fact that τ_i has a density, one obtains

$$\begin{aligned}
\gamma_A^i(t,\sigma_j) - \gamma_A^i(t_0,\sigma_j) &\geq se^{-rt}\mathbf{P}(\tau_i \geq t, \theta_j \geq t) - se^{-rt_0}\mathbf{P}(\tau_i \geq t_0, \theta_j > t_0) && (17)\\
&\quad + s\mathbf{E}\left[e^{-r\theta_j}1_{t_0<\theta_j<\min(\tau_i,t)}\right] + \gamma\mathbf{E}\left[e^{-r\tau_i}1_{t_0\leq\tau_i<\min(\theta_j,t)}\right]. && (18)\\
&\geq se^{-rt}\mathbf{P}(\tau_i \geq t, \theta_j > t_0) - se^{-rt_0}\mathbf{P}(\tau_i \geq t_0, \theta_j > t_0) + \gamma\mathbf{E}\left[e^{-r\tau_i}1_{t_0\leq\tau_i<\min(\theta_j,t)}\right] && (19)
\end{aligned}$$

Using **Fact 3b**, this implies that $\displaystyle\frac{1}{e^{-rt_0}\mathbf{P}(\tau_i \geq t_0, \theta_j > t_0)} \liminf_{t\searrow t_0} \frac{\gamma_A^i(t,\sigma_j) - \gamma_A^i(t_0,\sigma_j)}{t-t_0}$ is at least equal to

$$\lim_{t\searrow t_0} \frac{1}{t-t_0}\left\{se^{-r(t-t_0)}\mathbf{P}(\tau_i \geq t \mid \tau_i \geq t_0, \theta_j > t_0) + \gamma\mathbf{E}\left[e^{-r(\tau_i-t_0)}1_{\tau_i<t} \mid \tau_i \geq t_0, \theta_j > t_0\right] - s\right\},$$

which is equal to $W'(\xi_i(t_0+))$. This concludes the proof of **Fact 5**.

Since the inequality $\xi_i(t_0+) > p_*$ implies that $\xi_i(t+) > p_*$ for all $t > t_0$ close enough to t_0, the next fact is a direct consequence of **Fact 5**.

Fact 6 Assume that $\xi_i(t_0+) > p_*$. Then $t \mapsto \gamma_A^i(t,\sigma_j)$ is increasing over $[t_0, t_0+\varepsilon]$, for $\varepsilon > 0$ small enough.

B.2 A priori best-reply properties

In this section, we let σ_j be an arbitrary strategy of player j, and we let σ_i be a best reply to σ_j. We denote by S_i and S_j the supports of σ_i and σ_j.

We derive two properties which will help us pinning down the equilibrium set. In Lemmas 1 and 2, we first give a formal content to the intuitive assertion that a player will not drop as long as his belief exceeds p_*. Since the belief ξ_i need not be continuous, this will surprisingly turn out to be somewhat delicate. Next, in Lemma 3, we prove that if there is there is a positive probability that player j will drop at t_0, then it cannot be optimal for player i to drop out just before t_0.

Lemma 1 *One has*

- $\xi_i(t+) \leq p_*$ *for all* $t \in S_i$.
- $\xi_i(t)(=\xi_i(t+)) = p_*$ *for all* $t \in S_i \setminus S_j$.

Proof. We start with the first statement. Recall that, by the best-reply property, one has $\gamma_A^i(t,\sigma_j) = \gamma_A^i(\sigma_i,\sigma_j)$ for σ_i-a.e. $t \in S_i$.

We argue by contradiction, and assume that $\xi_i(t_0+) > p_*$ for some $t_0 \in S_i$. By **Fact 6**, there is $\bar{t} > t_0$ such that $\gamma_A^i(\bar{t}, \sigma_j) > \gamma_A^i(t, \sigma_j)$ for all $t \in [t_0, \bar{t})$, hence $\sigma_i([t_0, \bar{t})) = 0$. On the other hand, since $\gamma_A^i(t_0+, \sigma_j) \geq \gamma_A^i(t_0, \sigma_j)$ and since $\gamma_A^i(\cdot, \sigma_j)$ is left-continuous, one also has $\gamma_A^i(t, \sigma_j) < \gamma_A^i(\bar{t}, \sigma_j)$ for all $t \in [\underline{t}, t_0]$, as soon as $\underline{t} < t_0$ is close enough to t_0. Therefore, $\sigma_i([\underline{t}, t_0]) = 0$, and thus $\sigma_i([\underline{t}, \bar{t})) = 0$. But this contradicts the assumption that $t_0 \in S_i$.

We turn to the second statement. Let I be an open interval containing t, and such that $I \cap S_j = \emptyset$. In particular, $x \mapsto \gamma_A^i(x, \sigma_j)$ is continuous on I. Thus,

$$\gamma_A^i(x, \sigma_j) \leq \gamma_A^i(t, \sigma_j) = \gamma_A^i(\sigma_i, \sigma_j),$$

for each $x \in I$. The result is then a direct consequence of **Fact 4**. ■

Lemma 2 *Let $t_0 \in \mathbf{R}_+$. If $\xi_i(t_0) > p_*$, then $\sigma_i([t_0 - \varepsilon, t_0]) = 0$ for each $\varepsilon > 0$ small enough.*

Proof. Since ξ_i is left-continuous, one has $\xi_i(t) > p_*$ for all $t \in [\underline{t}, t_0]$, as soon as $\underline{t} < t_0$ is close enough to t_0. By **Fact 6**, this implies that $t \mapsto \gamma_A^i(t, \sigma_j)$ is increasing on the interval $[\underline{t}, t_0]$. Hence, $\sigma_i([\underline{t}, t_0)) = 0$.

We next discuss two cases. If $\sigma_j(\{t_0\}) = 0$, then $\xi_i(t_0+) = \xi_i(t_0) > p_*$. Using the proof of **Fact 6**, the map $t \mapsto \gamma_A^i(t, \sigma_j)$ is increasing on the interval $[t_0, \bar{t}]$, as soon as $\bar{t} > t_0$ is close enough to t_0. It follows in that case that $\sigma_i([t_0, \bar{t})) = 0$, as desired.

If instead $\sigma_j(\{t_0\}) > 0$, then $\xi_i(\cdot)$ is discontinuous at t_0. Since $\xi_i(t_0)$ is a weighted average of $\phi(t_0)$ and of $\xi_i(t_0+)$, one has either $\phi(t_0) > p_*$ or $\xi_i(t_0+) > p_*$. In the former case, one has $W(\phi(t_0)) > s$, and thus $\gamma_A^i(t_0+, \sigma_j) > \gamma_A^i(t_0, \sigma_j)$, using (14). It follows that $\sigma_i(\{t_0\}) = 0$. In the latter case, by **Fact 6** and as in the preceding paragraph, one has $\sigma_i([t_0, \bar{t})) = 0$ for $\bar{t} > t_0$ close enough to t_0. ■

We stress that Lemmas 1 and 2 do not rule out the possibility that $\xi_i(t) > p_*$ for some $t \in S_i$. Indeed, Lemmas 1 and 2 are consistent with a situation in which $\xi_i(t_0) > p_* \geq \xi_i(t_0+)$, and a strategy σ_i that assigns probability zero to the interval $[\underline{t}, t_0]$, yet assigns a positive probability to the interval $(t_0, t_0 + \varepsilon)$, for all $\varepsilon > 0$.

While the conclusion that $\xi_i(t) \leq p_*$ for all $t \in S_i$ is nevertheless valid, it will only be established as a corollary to our characterization results, and we do not know of any simple and direct proof.

Lemma 3 *Let $t_0 \in \mathbf{R}_+$ be given. Assume that $\sigma_j(\{t_0\}) > 0$. Then $\sigma_i([t_0 - \varepsilon, t_0]) = 0$ for some $\varepsilon > 0$.*

Proof. We handle first a few easy cases. If $\gamma_A^i(t_0+, \sigma_j) > \gamma_A^i(t_0, \sigma_j)$, then by left-continuity, one has $\gamma_A^i(t, \sigma_j) < \gamma_A^i(\bar{t}, \sigma_j)$ for all $t, \bar{t}$ close enough to t_0 and such that $t \leq t_0 < \bar{t}$. The result follows in that case.

If now $\xi_i(t_0) > p_*$, the conclusion follows by Lemma 2.

We are thus left with the case in which $\xi_i(t_0) \leq p_*$ and $\gamma_A^i(t_0+, \sigma_j) = \gamma_A^i(t_0, \sigma_j)$. Using equation (14), one must have $W(\phi(t_0)) = s$, so that $\phi(t_0) \leq p_*$. By Lemma 1, one has $\xi_i(t_0+) \leq p_*$ as well. Since ξ_i is discontinuous at t_0, and since $\xi_i(t_0)$ is a weighted average of $\xi_i(t_0+)$ and of $\phi(t_0)$, this implies that $\xi_i(t_0) < p_*$.

We claim that one must then have

$$\limsup_{t \nearrow t_0} \frac{\gamma_A^i(t_0, \sigma_j) - \gamma_A^i(t, \sigma_j)}{t_0 - t} < 0. \tag{20}$$

Since $\gamma_A^i(t_0, \sigma_j) = \gamma_A^i(t_0+, \sigma_j) = \gamma_A^i(\sigma_i, \sigma_j)$, the conclusion will follow.

We rely on equation (12). Since $t \mapsto W(\phi(t))$ is a C^1 function, with $W(\phi(t_0)) = s$, and by **Facts 3** and **3b**, the difference between $\dfrac{1}{e^{-rt}\mathbf{P}(\tau_i \geq t, \theta_j \geq t)} \times \dfrac{\gamma_A^i(t_0, \sigma_j) - \gamma_A^i(t, \sigma_j)}{t_0 - t}$ and

$$\frac{1}{t_0 - t}\left\{ s e^{-r(t_0 - t)} \mathbf{P}(\tau_i \geq t_0 \mid \tau_i \geq t, \theta_j \geq t) + \gamma \mathbf{E}\left[e^{-r(\tau_i - t)} 1_{t \leq \tau_i < t_0} \mid \tau_i \geq t, \theta_j \geq t \right] - s \right\} \tag{21}$$

converges to zero. Observe that the expression in (21) is equal to the incremental payoff over s (normalized by $t_0 - t$) in $\mathcal{P}$ when using the policy $t_0 - t$, and when starting from a belief of $\xi_i(t)$. Since $\xi_i(t_0) < p_*$, the limit of this expression is therefore negative. This proves (20). ■

B.3 On non-atomic equilibria

Again, we let a strategy σ_j of player j be given. We set $F_j(t) := \sigma_j([0, t])$, which is also equal to $\mathbf{P}(\theta_j \leq t \mid R_j = B)$. We also introduce $H_j(t) := \displaystyle\int_0^t e^{-\lambda x}(1 - F_j(x))dx$. In this section, we prove the next result.

Proposition 5 *Let I be an interval in $\mathbf{R}^+$ with non-empty interior, such that $\sigma^j(\{t\}) = 0$ for each $t \in I$. Then the following two statements are equivalent:*

- *On the interval I, the map $t \mapsto \gamma_A^i(t, \sigma^j)$ is constant;*
- *On the interval I, the map H_j is C^2 and is a solution to the linear, second-order differential equation*

$$\frac{W(\phi(t)) - s}{\lambda(\gamma - s)} \Big(H''(t) + \lambda H'(t) \Big) = (\phi(t) - p_*)H'(t) + \lambda \phi(t)(\psi(t) - p_*)H(t). \tag{22}$$

Proof. We prove that the first statement implies the second one. Fix $x \in I$. For $y > x$ in I, the equality $\gamma^i_A(x, \sigma_j) = \gamma^i_A(y, \sigma_j)$ writes

$$\begin{aligned} 0 &= \frac{1}{y-x}\left(s\left\{e^{-r(y-x)}\mathbf{P}(\theta_j \geq y, \tau_i \geq y \mid \min(\theta_j, \tau_i) \geq x) - 1\right\}\right. \\ &\quad + \mathbf{E}\left[W(\phi(\theta_j))e^{-r(\theta_j - x)}1_{\theta_j < \min(y, \tau_i)} \mid \min(\theta_j, \tau_i) \geq x\right] \\ &\quad \left. + \gamma\mathbf{E}\left[e^{-r(\tau_i - x)}1_{\tau_i < \min(y, \theta_j)} \mid \min(\theta_j, \tau_i) \geq x\right]\right). \end{aligned}$$

Taking the limit $y \searrow x$ in the previous equality, since $t \mapsto W(\phi(t))e^{-r(t-x)}$ is C^1, and by **Facts 1**, **2**, **3** and **3b**, one obtains

$$0 = -rs + \lambda\xi_i(x)(\gamma - s) + (W(\phi(x)) - s)\lim_{y \to x}\frac{\mathbf{P}(\theta_j < y \mid \min(\theta_j, \tau_i) \geq x)}{y - x}. \tag{23}$$

Write $N(x, y) := \mathbf{P}(\tau_i \geq x, \theta_j \in [x, y))$, and $D(x) := \mathbf{P}(\tau_i \geq x, \theta_j \geq x)$. With these notations, equation (23) rewrites

$$(W(\phi(x)) - s)\lim_{y \searrow x}\frac{N(x, y)}{y - x} = rsD(x) - \lambda(\gamma - s)\mathbf{P}(R^i = G^i, \min(\tau_i, \theta_j) \geq x). \tag{24}$$

We introduce the function $\tilde{F}_j(t) := \mathbf{P}(\theta_j \leq t \mid R_j = G)$, which is related to F_j by the identity

$$1 - \tilde{F}_j(t) = e^{-\lambda t}(1 - F_j(t)) + \int_0^t \lambda e^{-\lambda z}(1 - F_j(z))dz.$$

With these notations at hand, one has

$$N(x, y) = \left(q_{GG}e^{-\lambda x} + q_{GB}\right)(\tilde{F}_j(y) - \tilde{F}_j(x)) + \left(q_{GB}e^{-\lambda x} + q_{BB}\right)(F_j(y) - F_j(x)).$$

Using the relation between $\tilde{F}_j$ and F_j, and the continuity of F_j, equation (24) implies that F_j is C^1 on the interval I, and therefore, $\tilde{F}_j$ is C^1 as well, hence H_j is C^2. In addition, $\tilde{F}'_j(t) = e^{-\lambda t}F'_j(t)$ for each $t \in I$. It follows that

$$\frac{\partial N}{\partial y}(x, x) = (q_{GG}e^{-2\lambda x} + 2q_{GB}e^{-\lambda x} + q_{BB})F'_j(x). \tag{25}$$

On the other hand, $D(x)$ is given by

$$D(x) = q_{GG}e^{-\lambda x}(1 - \tilde{F}_j(x)) + q_{GB}e^{-\lambda x}(1 - F_j(x)) + q_{GB}(1 - \tilde{F}_j(x)) + q_{BB}(1 - F_j(x)) \tag{26}$$

while

$$\mathbf{P}(R_i = G_i, \min(\tau_i, \theta_j) \geq x) = q_{GG}e^{-\lambda x}(1 - \tilde{F}_j(x) + q_{GB}e^{-\lambda x}(1 - F_j(x)). \tag{27}$$

When plugging (25), (26) and (27) into (24), when using the relation between $\tilde{F}_j$ and F_j, and after obvious algebraic manipulations, one obtains

$$
\begin{aligned}
& (W(\phi(x)) - s)\left(q_{GG}e^{-2\lambda x} + 2q_{GB}e^{-\lambda x} + q_{BB}\right) F_j'(x) && (28)\\
= \; & (1 - F_j(x))\left\{rs\left(q_{GG}e^{-2\lambda x} + 2q_{GB}e^{-\lambda x} + q_{BB}\right) - \lambda(\gamma - s)(q_{GG}e^{-2\lambda x} + q_{GB}e^{-\lambda x})\right\} && (29)\\
& + \left(\int_0^x \lambda e^{-\lambda t}(1 - F_j(t))dt\right) \times \left(rs(q_{GG}e^{-\lambda x} + q_{GB}) - \lambda(\gamma - s)q_{GG}e^{-\lambda x}\right). && (30)
\end{aligned}
$$

Observe next that $q_{GG}e^{-2\lambda x} + 2q_{GB}e^{-\lambda x} + q_{BB} = \mathbf{P}(\tau_i \geq x, \tau_j \geq x)$, while

$$q_{GG}e^{-2\lambda x} + q_{GB}e^{-\lambda x} = \mathbf{P}(R_i = G, \tau_i \geq x, \tau_j \geq x) = \phi(x)\mathbf{P}(\tau_i \geq x, \tau_j \geq x).$$

On the other hand,

$$\frac{q_{GG}e^{-\lambda x}}{q_{GG}e^{-\lambda x} + q_{GB}} = \frac{\mathbf{P}(R_i = G, R_j = G, \tau_i \geq x)}{\mathbf{P}(R_j = G, \tau_i \geq x)} = \psi(x).$$

Therefore, one has

$$
\begin{aligned}
(W(\phi(x)) - s)\, F_j'(x) \; = \; & (1 - F_j(x))\,(rs - \lambda(\gamma - s)\phi(x)) \\
& + \phi(x)e^{\lambda x}\,(rs - \lambda(\gamma - s)\psi(x)) \int_0^x \lambda e^{-\lambda t}(1 - F_j(t))dt
\end{aligned}
$$

or equivalently, using $p_* := \frac{rs}{\lambda(\gamma - s)}$,

$$\frac{W(\phi(x)) - s}{\lambda(\gamma - s)} F_j'(x) = (1 - F_j(x))(p_* - \phi(x)) + \lambda\phi(x)(p_* - \psi(x))e^{\lambda x} \int_0^x e^{-\lambda t}(1 - F_j(t))dt.$$

Since $H_j(x) = \displaystyle\int_0^t e^{-\lambda t}(1 - F_j(t))dt$, it follows that H_j solves the equation

$$\frac{W(\phi(x)) - s}{\lambda(\gamma - s)}\left(H_j''(t) + \lambda H_j'(t)\right) = (\phi(x) - p_*)H_j'(t) + \lambda\phi(x)(\psi(x) - p_*)H_j(t)dt$$

on the interval I, as desired.

Conversely, if $H_j(x) := \displaystyle\int_0^x e^{-\lambda t}(1 - F_j(t))$ is a C^2 function and solves (23) on the interval, then the above computation shows that the map $t \mapsto \gamma_A^i(t, \sigma_j)$ is differentiable on I, with a derivative equal to zero throughout I. It follows that it is a constant map, as desired. ■

B.4 Proof of Theorem 1 and of Proposition 2

We here specialize to the case $\rho > 0$. We start with the proof of Proposition 2.

B.4.1 Proof of Proposition 2

We let (σ_1, σ_2) be an equilibrium, and let $t_0 \in S_i$. Assume for a moment that $\xi_i(t_0) > p_*$. Since $\rho > 0$, this implies that $\xi_i(t+) > p_*$ throughout some neighborhood of t_0. Hence, by **Fact 6**, the map $t \mapsto \gamma_A^i(t, \sigma_j)$ is increasing over some neighborhood, which is in contradiction with $t_0 \in S_i$. Therefore, $\xi_i(t_0) \leq p_*$. Since $\rho > 0$, this implies $\phi(t_0) \leq p_*$. By Equation (14), this implies that $t \mapsto \gamma_A^i(\cdot, \sigma_j)$ is continuous at t_0, so that $\gamma_A^i(t_0, \sigma_j) = \gamma_A^i(\sigma_i, \sigma_j)$. In that case, $\gamma_A^i(t_0, \sigma_j) = \gamma_A^i(\sigma_i, \sigma_j)$ and $\xi_i(t_0) = \xi_i(t_0+)$.

We next claim that $\xi_i(t_0) = p_*$. Assume by way of contradiction that $\xi_i(t_0) < p_*$. We claim that $\gamma_A^i(t_0, \sigma_j) < \gamma_A^i(t, \sigma_j)$, provided $t < t_0$ is close enough to t_0. This will yield the desired contradiction. Since $\phi(t_0) < p_*$, the quantity $\dfrac{1}{e^{-rt}\mathbf{P}(\theta_j \geq t, \tau_i \geq t)} \times \dfrac{\gamma_A^i(t_0, \sigma_j) - \gamma_A^i(t, \sigma_j)}{t - t_0}$ is equal to

$$\frac{1}{t_0 - t}\left\{se^{-r(t_0-t)}\mathbf{P}(\tau_i \geq t_0, \theta_j \geq t_0 \mid \min(\tau_i, \theta_j) \geq t) - s\right.$$
$$\left. + s\mathbf{E}\left[e^{-r(\theta_j - t)}1_{t \leq \theta_j < \min(\tau_i, t_0)} \mid \min(\tau_i, \theta_j) \geq t\right] + \gamma\mathbf{E}\left[e^{-r(\tau_i - t)}1_{t \leq \tau_i < \min(\theta_j, t_0)} \mid \min(\tau_i, \theta_j) \geq t\right]\right\}.$$

Since $x \mapsto e^{-r(x-t)}$ is C^1, and by **Facts 1** through **3**, this expression has the same limit as

$$\frac{1}{t_0 - t}\left\{se^{-r(t_0-t)}\mathbf{P}(\tau_i \geq t_0 \mid \min(\tau_i, \theta_j) \geq t) + \gamma\mathbf{E}\left[e^{-r(\tau_i - t)}1_{\tau_i < t_0} \mid \min(\tau_i, \theta_j) \geq t\right] - s\right\}.$$

The latter expression is equal to the difference between the payoff in $\mathcal{P}$ induced by the policy $t_0 - t$, when starting with a belief of $\xi_i(t)$, and s, divided by $t_0 - t$. Since $\xi_i(t_0) < 0$, the limit is therefore negative, as desired.

Therefore, $\xi_i(t_0) = p_*$. We next prove that $\sigma_j(\{t_0\}) = 0$, so that $\xi_i(t_0+) = \xi_i(t) = p_*$, as desired. Assume to the contrary that $\sigma_j(\{t_0\}) > 0$. By Lemma 3, one has $\sigma_i([t_0 - \varepsilon, t_0]) = 0$ for some $\varepsilon > 0$. On the other hand, and since ξ_i is discontinuous, one has $\xi_i(t_0+) > p_*$. Using **Fact 6**, this implies that $\sigma_i([t_0, t_0 + \varepsilon]) = 0$ for $\varepsilon > 0$ small enough. Hence σ_i assigns probability zero to some neighborhood of t_0. But this is in contradiction with $t_0 \in S_i$.

B.4.2 Proof of Theorem 1

The proof is organized as follows. We first assume the existence of a symmetric equilibrium and show uniqueness. We next prove existence.

Let (σ, σ) be a symmetric equilibrium, with support S.

Lemma 4 *The set S is equal to $[T_p, T_\psi]$.*

Proof. Since (σ,σ) is an equilibrium and by Lemma 3, the distribution σ has no atom, and thus $\xi_1(\cdot)=\xi_2(\cdot)$ is continuous on S. By Proposition 3, $\xi_1=\xi_2$ is equal to p_* on S. Finally, if I is any interval such that $\sigma(I)=0$, then ξ_i is decreasing on I. This implies that S is an interval, say $[T_1,T_2]$. Since $\sigma([0,T_1])=0$, and $\sigma([0,T_2])=1$, one has respectively $\xi_i(T_1)=p(T_1)$ and $\xi_i(T_2)=\psi(T_2)$, hence $T_1=T_p$, and $T_2=T_\psi$. ■

Define $F(t):=\sigma([0,t]$, and set $H(t):=\int_0^t e^{-\lambda x}(1-F(t))dt$. By Lemma 4, $H(T_p)=\int_0^{T_p} e^{-\lambda t}dt=\frac{1-e^{-\lambda T_p}}{\lambda}$. Since σ has no atom, the payoff function $t\mapsto\gamma_A^i(t,\sigma)$ is constant on the interval $[T_p,T_\psi]$. Therefore, by Proposition 5, H is C^2 on the interval $[T_p,T_\psi]$ and solves the equation (22) on that interval. Since $\rho>0$, one has $\phi(t)<p_*$, and therefore $W(\phi(t))=s$, for each $t\in[T_p,T_\psi]$. Hence the equation (22) boils down to the first-order equation

$$H'(x)=\frac{\lambda\phi(x)(\psi(x)-p_*)}{p_*-\phi(x)}H(x). \tag{31}$$

This equation has a unique solution such that $H(T_p)=\frac{1-e^{-\lambda T_p}}{\lambda}$. This shows that a symmetric equilibrium, if it exists, must be unique.

Conversely, consider the (unique) solution H to the equation (31) on the interval $[T_p,T_\psi]$, that satisfies the initial condition $H(T_p)=\frac{1-e^{-\lambda T_p}}{\lambda}$, and observe that H is smooth on the interval $[T_p,T_\psi]$. Set $F(t)=1-e^{\lambda t}H'(t)$ for each $t\in[T_p,T_\psi]$. We now prove that F is the cumulative of some mixed strategy, which forms a symmetric equilibrium.

Observe first that, using (31) and the boundary condition, one has $F(T_p)=0$, and $F(T_\psi)=1$. It can be checked that the derivative of F is of the sign of ρ on the interval $[T_p,T_\psi]$, so that F is increasing.[21]

Let σ be the probability distribution defined by $\sigma([0,t])=F(t)$. Note that σ has no atom. When facing the strategy σ, the belief of player i solves $\xi_i(t)=\mathbf{P}(R_i=G\mid\tau_i\geq t)>p_*$ for $t<T_p$, and $\xi_i(t)=\psi(t)<p_*$ for $t>T_\psi$. Hence any best-reply of player i to σ assigns probability 1 to $[T_p,T_\psi]$. By Proposition 5, the map $t\mapsto\gamma_A^i(t,\sigma)$ is constant on $[T_p,T_\psi]$. Hence, any strategy that assigns probability 1 to $[T_p,T_\psi]$ is a best-reply to σ. Therefore, (σ,σ) is an equilibrium.

B.5 Proof of Proposition 3

We assume here that the two arms are negatively correlated, $\rho<0$. We prove that the game has a unique equilibrium, which happens to be symmetric. We first prove uniqueness, then existence, of an equilibrium.

[21]Details of the computation are available upon request.

Thus, assume first that the game has an equilibrium, say (σ_1, σ_2). Note that since $\rho < 0$, the belief $\xi_i(\cdot)$ is decreasing on $\mathbf{R}_+$. We denote by S_i the support of the distribution σ_i.

Lemma 5 *One has $S_1 = S_2 = [T_p, \hat{T}]$, for some $\hat{T}$ such that $T_p < \hat{T} < T_\phi f$.*

Proof. We first argue by way of contradiction that $S_1 = S_2$, and thus assume that there exists $t_0 \in S_1 \setminus S_2$. By Lemma 1, one has $\xi_i(t_0) = p_*$. By **Fact 6**, and since ξ_1 is decreasing, one has $t_0 = \min S_1$. Hence, ξ_2 is simply equal to $\mathbf{P}(R_2 = G \mid \tau_2 \geq t)$ for $t < t_0$. This implies that $\xi_2(t) \geq \xi_1(t)$, and thus, $\xi_2(t) > p_*$, for each $t < t_0$. By **Fact 6**, one must then have $t_0 < \min S_2$.

Set $\underline{t} := \inf\{t > t_0 : t \in S_1 \cup S_2\}$. Observe that $\xi_1(t) < p_*$ for each $t \in (t_0, \min S_2)$, hence $t \notin S_1$ by Lemma 1. This implies that $\underline{t} = \min S_2$. We claim that $\underline{t}$ belongs to S_1 as well. If instead $\underline{t} \notin S_1$. then $\xi_2(\underline{t}) = p_*$ by Lemma 1. On the other hand, and since $\underline{t} = \min S_2$, one has $\xi_1(t) \geq \xi_2(t)$ for $t \leq \underline{t}$, which yields $\xi_2(\underline{t}) < p_*$ since $\xi_1(t) < p_*$ for each $t > t_0$ – a contradiction, and thus $\underline{t} \in S_1 \cap S_2$, as claimed. Note that we have proven as well that $\xi_i(\underline{t}) < p_*$ for $i = 1, 2$.

By Lemma 3, at least one of the two strategies σ_1, σ_2 does not have an atom at $\underline{t}$. Fix $i \in \{1, 2\}$ such that $\sigma_i(\{\underline{t}\}) = 0$. Then, $\gamma_A^j(\cdot, \sigma_i)$ is continuous at $\underline{t}$, so that $\gamma_A^j(\underline{t}, \sigma_i) = \gamma_A^j(\sigma_i, \sigma_j)$. By **Fact 4**, the map $t \mapsto \gamma_A^i(t, \sigma_j)$ is decreasing over the interval $[t_0, \underline{t}]$. This is in contradiction with the equality $\gamma_A^j(\underline{t}, \sigma_i) = \gamma_A^j(\sigma_i, \sigma_j)$. This concludes the proof that $S_1 = S_2$.

We next prove, again by way of contradiction, that the common set $S_1 = S_2$ is an interval. Assume instead that there are two dates $\underline{t} < \bar{t}$ in S, such that $(\underline{t}, \bar{t}) \cap S = \emptyset$.

Wlog, using Lemma 3, assume that $\sigma_2(\{\underline{t}\}) = 0$, so that $\gamma_A^2(\underline{t}, \sigma_1) = \gamma_A^2(\sigma_1, \sigma_2)$. Since $\underline{t} \in S$, one has $\sigma_1(\{\underline{t}\}) = \sigma_2(\{\underline{t}\}) = 0$,[22] one has $\xi_2(\underline{t}) \leq p_*$. By **Fact 4** again, this implies that the map $t \mapsto \gamma_A^2(t, \sigma^1)$ is decreasing over the interval $(\underline{t}, \bar{t}]$ – a contradiction. ■

By Lemma 3, none of the two distributions σ_1 and σ_2 can possibly have an atom in $(T_p, \hat{T}]$, and at most one of the two distributions has an atom at T_p. Assume that, say, that σ_1 has no atom. We denote by F_1 the cdf of σ_1, and introduce $H_1(t) = \displaystyle\int_0^t e^{-\lambda x}(1 - F_1(x))dx$. Since σ_1 is non-atomic, the map $t \mapsto \gamma_A^2(t, \sigma_1)$ is constant on the interval $[T_p, \hat{T}]$. By Proposition 5, H_1 is C^2 on the interval $[T_p, \hat{T}]$, and solves the linear equation (22) on this interval. Since $W(\phi(t)) > s$ on the interval $[T_p, T_\phi)$, the coefficient of H_1'' does not vanish on this interval, and there is a unique solution to (22) on the interval $[T_p, T_\phi)$ that assumes given initial values for $H_1(T_p)$ and $H_1'(T_p)$.

On the other hand, note that H_1 satisfies the boundary condition $H_1(T_p) = \frac{1-e^{-\lambda T_p}}{\lambda}$, $H_1'(T_p) = e^{-\lambda T_p}$. Hence, H_1, and therefore F_1 is uniquely defined. This in turn defines uniquely $\hat{T}$ as the smallest t such that $F_1(t) = 1$.

[22]Otherwise, by Lemma 3, and since $S_1 = S_2$, $\underline{t}$ would be a common atom of σ_1 and of σ_2 – contradicting Lemma 3.

We next proceed to show that $\sigma_1 = \sigma_2$. We denote by F_2 the cdf of σ_2, and introduce the map $H_2(t) := \int_0^t e^{-\lambda x}(1 - F_2(x))dx$, for $t > T_p$. We argue by contradiction and assume that $F_2(T_p) > 0 (= F_1(T_p))$. Since σ_2 has no atom in $(T_p, \hat{T}]$, one has $\gamma_A^1(t, \sigma_2) = \gamma_A^1(\sigma_1, \sigma_2)$ for each $t \in (T_p, \hat{T}]$, and therefore, H_2 solves (22) on the interval $(T_p, \hat{T}]$. In addition, when setting $H_2(T_p) := \frac{1 - e^{-\lambda T_p}}{\lambda}$ and $H_2'(T_p) = e^{-\lambda T_p}(1 - F_2(T_p))$, the map H_2 is C^2 on the closed interval $[T_p, \hat{T}]$ and solves (22) on $[T_p, \hat{T}]$.

We set $\Omega(t) := (1 - F_1(t))H_2(t) - (1 - F_2(t))H_1(t)$. Note that $\Omega(T_p) > 0$, while $\Omega(\hat{T}) = 0$ since $F_1(\hat{T}) = F_2(\hat{T}) = 1$.

Set $T := \min\{t \in [T_p, \hat{T}] : \Omega(t) = 0\}$. Since $\Omega(\hat{T}) = 0$, one has $T \in [T_p, \hat{T}]$ and $\Omega(T) = 0$. Since $\Omega(T_p) > 0$, one has $\Omega(t) > 0$ for each $t \in [T_p, T)$. From (22), one has

$$\frac{F_i'(t)}{H_i(t)} = \lambda\phi(x)(p_* - \psi(x)) + (p_* - \phi(x))\frac{H_i'(x)}{H_i(x)},$$

for $t \in [T_p, \hat{T}]$. Hence, for fixed t, the ratio $\frac{F_i'(t)}{H_i(t)}$ is decreasing w.r.t. the ratio $\frac{1 - F_i(t)}{H_i(t)}$. Since $\Omega(t) > 0$ for $t \in [T_p, T)$, it follows that $\frac{F_1'(t)}{H_1(t)} < \frac{F_2'(t)}{H_2(t)}$ for $t \in [T_p, T)$, and therefore, $\Omega(\cdot)$ is increasing over the interval $[T_p, \hat{T}]$. This contradicts the fact that $\Omega(T_p) > \Omega(\hat{T})$. This concludes the proof that there is at most one equilibrium.

We now turn to existence. Since $W(\phi(x)) > s$ for $x \in [T_p, T_\phi)$, there is a C^2 function H that solves (22) on the interval $[T_p, T_\phi)$ and such that $H(T_p) = \frac{1 - e^{-\lambda T_p}}{\lambda}$ and $H'(T_p) = e^{-\lambda T_p}$.

Lemma 6 *Let $T \in [T_p, T_\phi)$ be such that $H'(t) > 0$ on $[T_p, T]$. Then $H''(t) + \lambda H'(t) \leq 0$ for each $t \in [T_p, T]$.*

In particular, H' is decreasing over $[T_p, T]$.

Proof. Set $T_1 := \max\{t \in [T_p, T] : H''(t) + \lambda H'(t) \leq 0\}$. From (22), one can check that $H''(T_p) + \lambda H'(T_p) < 0$, hence $T_1 > T_p$. We claim that $T_1 = T$. Assume to the contrary that $T_1 < T$. One has $H''(t) \leq -\lambda H'(t) < 0$ for each $t_in[T_p, T]$, hence H' is decreasing on $[T_p, T_1]$, and $H''(T_1) <$. Hence there exists $T` \in (T_1, T]$ such that H' is decreasing on $[T_1, T']$.

We rewrite equation (22) as

$$\alpha(t)(H''(t) + \lambda H'(t)) = a(t)H'(t) + b(t)H(t),$$

with $a(t) := e^{\lambda t}(rs\mathbf{P}(\min(\tau_1, \tau_2) \geq t) - \lambda(\gamma - s)\mathbf{P}(R^i = G, \tau_j \geq t))$, $b(t) := \lambda e^{\lambda t}(rs\mathbf{P}(R^i = G, \tau_j \geq t) -$ and $\alpha(t) > 0$ for each $t \in [T_p, T_\phi]$. It can be shown that $a(\cdot)$ is positive and decreasing, while $b(\cdot)$

is negative and decreasing on $[T_p, T']$. On the other hand, H' is positive and decreasing, while H is positive and increasing on $[T_p, T']$.

Therefore,

$$a(t)H'(t) + b(t)H(t) < a(T_p)H'(T_p) + b(T_p)H(T_p) \tag{32}$$

for each $t \in (T_p, T']$, and an easy computation shows that the right-hand side of (32) is equal to zero. Hence $H'' + \lambda H'$ is negative on $[T_1, T']$ – in contradiction with the definition of T_1. ■

Lemma 7 *There exists $t < T_\phi$ such that $H'(t) = 0$.*

Proof. Assume instead that $H'(t) > 0$ for each $t \in [T_p, T_\phi)$. Integrating (22), there are $c_1, c_2 \in \mathbf{R}$, such that

$$H'(t) + \lambda H(t) = c_1 + c_2 \int_{T_p}^{t} \frac{\phi(x) - p_*}{W(\phi(x)) - s} H'(x)dx + \lambda \int_{T_p}^{t} \frac{\phi(x)(\psi(x) - p_*)}{W(\phi(x)) - s} H(x)dx. \tag{33}$$

By Lemma 6, H' is decreasing and positive on $[T_p, T_\phi]$, and is therefore bounded. Note also that H is positive, increasing and bounded.

Since the map $x \mapsto W(\phi(x)) - s$ is C^1 and vanishes for $t = T_\phi$, the first integral in (33) has a finite limit when $t \to T_\phi$, while the second one converges to $-\infty$, which implies $H'(t) \to -\infty$, a contradiction. ■

Define $\hat{T} := \min\{t \in [T_p, T_\phi] : H'(t) = 0\}$. Define F by $F(t) = 0$ for $t \leq T_p$, $F(t) = 1 - e^{\lambda t}H'(t)$ for $t \in [T_p, \hat{T}]$, and $F(t) = 1$ for $t > \hat{T}]$. The map F is increasing on $[T_p, \hat{T}]$ by Lemma 6, and is continuous over $\mathbf{R}_+$. It is therefore the cdf of a non-atomic measure σ, which is concentrated on $[T_p, \hat{T}]$. We now argue that (σ, σ) is a symmetric equilibrium.

Since σ is non-atomic, the map $t \mapsto \gamma_A^i(t, \sigma)$ is continuous over the compact set $\mathbf{R}_+ \cup \{+\infty\}$, and a best-reply therefore exists. When facing σ, one has $\xi_i(t) > p_*$ for $t < T_p$, and $\xi_i(t) < p_*$ for $t > \hat{T}$, hence the support of any best-reply map is included in $[T_p, \hat{T}]$. By Proposition 5, the map $t \mapsto \gamma_A^i(t, \sigma)$ is constant on $[T_p, \hat{T}]$, and thus, any strategy with support in $[T_p, \hat{T}]$ is a best-reply to σ.

B.6 Proof of Proposition 3

The proof of Proposition 3 may be the more delicate, and we will remain on the informal side, to save much on notation.[23] We let $K \in \mathbf{N}$ be given, and will prove the existence of an equilibrium with a support of cardinal K. Set $\Omega := \{(u_1, v_1, \ldots, u_K, v_K) \in \mathbf{R}^{2K}$ such that $T_p = u_1 \leq v_1 \leq \cdots \leq u_K \leq T_\psi\}$. We will focus on strategies with a finite support S. Rather than

[23] Formal details are available from the authors.

viewing such a strategy σ as a probability distribution over S, we equivalently describe it by the conditional probability of dropping at t, $\sigma(\{t\})/\sigma([0,t))$.

Given $\omega \in \Omega$, we define $S_1(\omega) = \{u_1, \ldots, u_K\}$ and $S_2(\omega) = \{v_1, \ldots, v_K\}$.

Given $\omega \in \Omega$, we define inductively $x_1(\omega), \ldots, x_{K-1}(\omega)$ as follows. We set $x_1(\omega) = 0$ if $v_1 = u_1 (= T_p)$. If $v_1 > u_1$, we define $x_1(\omega) \in [0,1]$ by the requirement that, if the strategy σ_1 of player has a support in $S_1(\omega)$ and drops with probability $x_1(\omega)$ at u_1, then the belief $\xi_2(v_1)$ held by player 2 at v_1 is equal to p_*. Similarly, $x_k(\omega)$ is set to 0 if $v_k = u_k$, and otherwise defined by the condition that $\xi_2(v_k) = p_*$, if σ_1 has a support in $S_1(\omega)$ and drops with probability $x_l(\omega)$ at time u_l $(l \leq k)$.

Somewhat symmetrically, the values $y_l(\omega)$ $(l = 1, \ldots, K-1)$ are defined inductively in such a way that the belief $\xi_1(u_k)$ $(k > 1)$ held by player 1 is equal to p_* if σ_2 drops with probability $y_l(\omega)$ at dates v_l, $l = 1, \ldots, k-1$.

Whether or not all dates in ω are distinct, one then has $\xi_1(u_k) = \xi_2(v_k) = p_*$, for each $k \in \{1, \ldots, K\}$. In addition, $x_k(\omega)$ and $y_k(\omega)$ lie in $[0,1]$ and are continuous in ω.

We introduce some additional piece of notation. Given $\omega \in \Omega$, and $\tilde{u}_k \in [v_{k-1}, v_k]$, we denote by $\omega|\tilde{u}_k$ the point in Ω obtained when substituting $\tilde{u}_k$ to u_k.

We next proceed by defining a continuous map from Ω into itself. It will be obtained as the composition of several maps $\alpha_k(\omega)$ $(k > 1)$ and $\beta_k(\omega)$ $(k < K)$.

We start with $\alpha_k(\omega)$. Let $\omega \in \Omega$ be given. If $v_k = v_{k-1}$, we set $\alpha_k(\omega) = v_k$. If instead $v_k > v_{k-1}$, and for any $\tilde{u}_k \in (v_{k-1}, v_k)$, we denote by $\Delta_k^2(\tilde{u}_k)$ the difference in player 2's expected payoffs induced by the strategies v_{k-1} and v_k, when facing the strategy of player 1 which drops out with probability $x_l(\omega|\tilde{u}_k)$ at time $u_1, \ldots, u_{k-1}, \tilde{u}_k, \ldots$. We claim that $\Delta_k^2(\cdot)$ is decreasing over the open interval (v_{k-1}, v_k). To see this, let $\underline{u}_k < \bar{u}_k$ be given in (v_{k-1}, v_k). We denote by $\sigma_1^{\underline{u}_k}$ and $\sigma_1^{\bar{u}_k}$ the corresponding two strategies of player 1.

Using the strategy v_k rather than v_{k-1} , when facing either $\sigma_1^{\underline{u}_k}$ or $\sigma_1^{\bar{u}_k}$, may yield a different outcome only if $\min(\tau_2, \theta_1) \geq v_{k-1}$, and we henceforth assume that this event holds. The continuation payoff is then equal to s when using v_{k-1}. We will prove that the continuation payoff of strategy v_k is higher when facing $\sigma_1^{\underline{u}_k}$ than when facing $\sigma_1^{\bar{u}_k}$. We proceed in several steps.

Note first that to ensure a belief of p_* at time v_k, it must be that $x_k(\omega|\underline{u}_k) > x_k(\omega|\bar{u}_k)$. Next, everything else being kept constant, the expected payoff of player 2 (induced by the strategy v_k) is linear in the probability that player 1 drops out at $\underline{u}_k$ (or at $\bar{u}_k$). In addition, the higher this probability, the more informative if player 1's decision at time $\underline{u}_k$, and therefore, the higher this payoff.

This implies that it will be sufficient to prove the following statement. (Conditional on $\tau_2 > v_{k-1}, \theta_1 > v_{k-1}$), the expected payoff induced by the strategy v_k is higher when facing a strategy that stops with probability one at $\underline{u}_k$ (in the absence of payoffs) than when facing a strategy that stops with probability one at $\bar{u}_k$.

Note that the outcomes induced by these two profiles are identical, as soon as either $\tau_1 < \underline{u}_k$ or $\tau_2 < \underline{u}_k$. Hence, we may and will condition on the event $\min(\tau_1, \tau_2) \geq \underline{u}_k$. In that event, the continuation payoff induced by the first profile as of time $\underline{u}_k$ is equal to s, since player 2 drops out at $\underline{u}_k$, immediately following player 1. On the other hand, under the second profile, player 2 will either drop out at $\bar{u}_k$ if $\tau_1, \tau_2 \geq \bar{u}_k$, or remain active then, possibly forever. If $\bar{u}_k$ differs from $\underline{u}_k$ by an infinitesimal amount of time, and using the same arguments as in earlier proofs, this expected payoff is approximately equal to

$$s + \phi(t) \times \lambda dt \times (\gamma - s) + \phi(t) \times \lambda dt \times (W(\psi(t)) - s) - rsdt.$$

Using the fact that $\underline{u}_k \geq T_p$, so that $p(t) \leq p_*$, it is not difficult to show that this payoff does not exceed s, at least when ρ is sufficiently close to one, and the outside option s is close enough to zero.

This concludes the proof that Δ_k^2 is decreasing wrt $\tilde{u}_k$. Continuity of Δ_k^2 on the open interval (v_{k-1}, v_k) is easy to check. Finally, observe that, on the one hand, $\Delta_k^2(\cdot) > 0$ when $\tilde{u}_k$ is very close to v_{k-1} (the proof follows that of **Fact 1**), while $\Delta_k^2(\cdot) < 0$ when $\tilde{u}_k$ is very close to v_k. The reason there is that in the limit $\tilde{u}_k \to v_k$, player 2 will drop out at time v_k iff $\tau_2 \geq v_k$. Hence there is a unique value, $\hat{u}_k$, such that $\Delta_k^2(\hat{u}_k) = 0$, and we set $\alpha_k(\omega) = (\omega|\hat{u}_k)$.

The definition of $\beta_k(\omega)$ slightly differs, to reflect the disymmetry between the two players. Let $\omega \in \Omega$ be given. If $u_k = u_{k+1}$, we set $\beta_k(\omega) = u_k$. Otherwise, we set $\beta_k(\omega) = (\omega|\hat{v}_k)$, where $\hat{v}_k \in (u_k, u_{k+1})$ is the unique point such that the expected payoffs induced by the two strategies u_k and u_{k+1} against the strategy that drops out with probability $y_l(\omega|\hat{v}_k)$ at time $v_1, \ldots, v_{k-1}, \hat{v}_k, \ldots$.

The map $\alpha_K \circ \beta_K \circ \cdots \circ \alpha_1 \circ \beta_1$ maps continuously Ω into itself, hence has a fixed point, $\omega_* = (u_1^*, v_1^*, \ldots, u_K^*, v_K^*)$. It follows by construction that $T_p = u_1^* < v_1^* < \cdots u_K^* < v_K^* = T_\psi$, and that equilibrium requirements are met by the strategies that drop out respectively with probability $x_l(\omega_*)$ at time u_l^*, and with probability $y_l(\omega_*)$ at time v_l^*.